

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 01:17:11 ; Search time 12208 Seconds
(without alignments)
11919.563 Million cell updates/sec

Title: US-09-923-515-3_COPY1_5000
Perfect score: 5000
Sequence: 1 ctggagttggagacactt.....caagcttgcatctatgac 5000

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl1.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_pro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_other.*
33: em_hgt_mus.*
34: em_hgt_pln.*
35: em_hgt_rod.*
36: em_hgt_mam.*
37: em_hgt_vrt.*
38: em_sy.*
39: em_hgtgo_hum.*
40: em_hgtgo_mus.*
41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5000	100.0	13938	9	HSALIPOA	X06290 Human mRNA
2	2697.8	54.0	4510	9	MACAPOA	J04635 Rhesus monk
3	955.4	19.1	2041	9	HSU19517	U19517 Human (apoa
4	826	16.5	1712	9	HSU19518	U19518 Human (apoa
5	442.8	8.9	1649	9	AF029691	AF029691 Papio ham
6	396.6	7.9	435	9	S79621	S79621 apolloprot
7	277	5.5	194321	2	AC027197	AC027197 Homo sapi
8	242.4	4.8	2711	9	MACCEPSC	J04697 Rhesus monk
9	236.6	4.7	1784	9	HUMPMGM	K02922 Human plasm
10	236.6	4.7	2433	6	AX463622	AX463622 Sequence
11	236.6	4.7	2497	6	AR037325	AR037325 Sequence
12	236.6	4.7	2497	6	AR082437	AR082437 Sequence
13	236.6	4.7	2497	6	AR085163	AR085163 Sequence
14	236.6	4.7	2497	6	AX448883	AX448883 Sequence
15	236.6	4.7	2497	6	HUMPLASM	M74220 Human plasm
16	236.6	4.7	2679	6	I06212	I06212 Sequence 14
17	236.6	4.7	2732	9	HSPWGR	X05199 Human mRNA
18	236.6	4.7	2753	6	A22096	A22096 plasmidogen
19	236.6	4.7	2753	6	I45623	I45623 Sequence 1
20	235	4.7	2286	6	AR105749	AR105749 Sequence
21	235	4.7	2286	6	I14104	I14104 Sequence 18
22	216.6	4.3	1134	6	AX417488	AX417488 Sequence
23	216.2	4.3	1134	6	AX395661	AX395661 Sequence
24	207.6	4.2	2771	10	BC014773	BC014773 Mus muscu
25	206	4.1	2720	10	MUSP1GN	J04766 Mouse plasm
26	202.8	4.1	194321	2	AC027197	AC027197 Homo sapi
27	199	4.0	2770	4	BTPLASMIN	X79402 B. caurus mr
28	195.6	3.9	1092	6	AX417490	AX417490 Sequence
29	190.2	3.8	1077	4	AY069985	AY069985 Canis fam
30	187.2	3.7	2745	4	EU03171	U33171 Ertinaceus e
31	182.6	3.7	2737	10	RNO242649	AJ242649 Rattus no
32	182	3.6	98357	2	AL391871	AL391871 Homo sapi
33	182	3.6	135751	2	AL396089	AL396089 Human DNA
34	182	3.6	167031	2	AC040907	AC040907 Homo sapi
35	181.6	3.6	2832	4	AF012297	AF012297 Macropus
36	172.4	3.4	141287	2	AC093404	AC093404 Pan trogl
37	172.4	3.4	141287	2	AC093404	AC093404 Pan trogl
38	169.4	3.4	5533	9	HUMAP0AKIV	L14005 Homo sapien
39	166	3.3	171363	2	AC084863	AC084863 Papio cyn
40	165.4	3.3	964	11	HS81D8T	AL591278 H. sapiens
41	165.4	3.3	163599	9	HSU81D8T	AL109933 Human DNA
42	165	3.3	220	9	HUMPGRRRA	M86879 H. sapiens p
43	164.4	3.3	179528	2	AC084862	AC084862 Papio cyn
44	162	3.2	220	9	HUMPGRRRB	M86880 H. sapiens p
45	157.2	3.1	219	9	HUMPGRRRC	M86881 H. sapiens p

ALIGNMENTS

RESULT 1
LOCUS HSALIPOA 13938 bp mRNA linear PRI 30-MAR-1995
DEFINITION Human mRNA for apolipoprotein(a).
ACCESSION X06290 M17399 X06696
VERSION X06290.1 GI:28619
KEYWORDS apolipoprotein; apolipoprotein A; glycoprotein; lipoprotein;
plasmidogen; serine protease.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 13938)
AUTHORS McLean,T.W.
TITLE Direct Submission

QY 361 TCAGAGCAGAAAGGAGCTCCGTCGGCCTCCGACTGTTACCCCGGTTCCAAAGCCTAGAG 420
|||||
Db 361 TCAGAGCAGAAAGGAGCTCCGTCGGCCTCCGACTGTTACCCCGGTTCCAAAGCCTAGAG 420
QY 421 GCTCCTTCGGAACAGACCGACTGAGCAAAAGCCTGGGGTGCAGAGATGCTACCAATGGT 480
|||||
Db 421 GCTCCTTCGGAACAGACCGACTGAGCAAAAGCCTGGGGTGCAGAGATGCTACCAATGGT 480
QY 481 AATGACACAGATTATGAGAGCAGATCTCCACACTGTATACAGGAAGAACTGCGCAAGCT 540
|||||
Db 481 AATGACACAGATTATGAGAGCAGATCTCCACACTGTATACAGGAAGAACTGCGCAAGCT 540
QY 541 TGGTCACTATGACACACACTGCGATAGTGGACCCCGAATACTACCAATGCTGGC 600
|||||
Db 541 TGGTCACTATGACACACACTGCGATAGTGGACCCCGAATACTACCAATGCTGGC 600
QY 601 TTGATCATGAACTACTGACAGGAATCCAGATGCTGTGGAGCTCTTATTGTTATACGAGG 660
|||||
Db 601 TTGATCATGAACTACTGACAGGAATCCAGATGCTGTGGAGCTCTTATTGTTATACGAGG 660
QY 661 GATCCGGGTGTCAGTGGGAGTACTGCAACTGACGCATGCTCAGAGCCAGAAAGGAGCT 720
|||||
Db 661 GATCCGGGTGTCAGTGGGAGTACTGCAACTGACGCATGCTCAGAGCCAGAAAGGAGCT 720
QY 721 GCGGTGCGCCTTCGACTGTATACCCGGTTCCAAAGCCTAGAGGCTCTTCCGAACAAGCA 780
|||||
Db 721 GCGGTGCGCCTTCGACTGTATACCCGGTTCCAAAGCCTAGAGGCTCTTCCGAACAAGCA 780
QY 781 CCGACTGAGCAAAAGCCTGGGGTGCAGAGTCTACCATGTTAATGACAGAGTTATGCA 840
|||||
Db 781 CCGACTGAGCAAAAGCCTGGGGTGCAGAGTCTACCATGTTAATGACAGAGTTATGCA 840
QY 841 GGCACATATCTCCACTGCTGACAGAAAGAACTGCGAAGCTTTGGTCACTATGACACCA 900
|||||
Db 841 GGCACATATCTCCACTGCTGACAGAAAGAACTGCGAAGCTTTGGTCACTATGACACCA 900
QY 901 CACTGCGATAGTGCAGACCCAGATACTACCAAAATGCTGGCTTGATCATGAACTACTGC 960
|||||
Db 901 CACTGCGATAGTGCAGACCCAGATACTACCAAAATGCTGGCTTGATCATGAACTACTGC 960
QY 961 AGGAATCCAGATGCTGTGGCAGCTCTTATTGTTATACAGAGGATCCGGGTGCAGGTGG 1020
|||||
Db 961 AGGAATCCAGATGCTGTGGCAGCTCTTATTGTTATACAGAGGATCCGGGTGCAGGTGG 1020
QY 1021 GAGTACTGCAACCTGACGCAATGCTCAGACGCGAAGAGGAGCTGCGGTGCGCTCGACT 1080
|||||
Db 1021 GAGTACTGCAACCTGACGCAATGCTCAGACGCGAAGAGGAGCTGCGGTGCGCTCGACT 1080
QY 1081 GTTACCCGCGTTCCAAAGCCTAGAGGCTCTTCCGAACAAGCACCGCATGAGCAAAAGGCT 1140
|||||
Db 1081 GTTACCCGCGTTCCAAAGCCTAGAGGCTCTTCCGAACAAGCACCGCATGAGCAAAAGGCT 1140
QY 1141 GGGGTGCAGAGTGTCTACCATGTTATGACAGAGTTATGAGGACACATATCTCCAACT 1200
|||||
Db 1141 GGGGTGCAGAGTGTCTACCATGTTATGACAGAGTTATGAGGACACATATCTCCAACT 1200
QY 1201 GTTCACAGAAAGAACTGCGAAGCTTGATCATGATGACACCACTGCGATGTGGGAGC 1260
|||||
Db 1201 GTTCACAGAAAGAACTGCGAAGCTTGATCATGATGACACCACTGCGATGTGGGAGC 1260
QY 1261 CCAGAACTACTACCAAAATGCTGGCTTGATCATGATGACACCACTGCGATGTGGGAGC 1320
|||||
Db 1261 CCAGAACTACTACCAAAATGCTGGCTTGATCATGATGACACCACTGCGATGTGGGAGC 1320
QY 1321 GCAAGCTCTTATTGTTATACAGAGGATCCCGGTGTGAGTGGGAGTACTGCAACTGACG 1380
|||||
Db 1321 GCAAGCTCTTATTGTTATACAGAGGATCCCGGTGTGAGTGGGAGTACTGCAACTGACG 1380
QY 1381 CAATGCTCAGACGAGAAAGGAGCTGCGGTGCGGCTCCGACTGTTTACCCCGGTTCCAAAGC 1440
|||||
Db 1381 CAATGCTCAGACGAGAAAGGAGCTGCGGTGCGGCTCCGACTGTTTACCCCGGTTCCAAAGC 1440

QY 1441 CTAGAGCTCTCTCCGAACCAAGCACCCGACTGAGCAAAAGCCTGGGGTGCAGAGATGCTTAC 1500
|||||
Db 1441 CTAGAGCTCTCTCCGAACCAAGCACCCGACTGAGCAAAAGCCTGGGGTGCAGAGATGCTTAC 1500
QY 1501 CATGGTAATGACAGAGTTATCGAGGACATATCTCCACACTGTCTACAGGAAGAACTGC 1560
|||||
Db 1501 CATGGTAATGACAGAGTTATCGAGGACATATCTCCACACTGTCTACAGGAAGAACTGC 1560
QY 1561 CAAGCTTGATCATGATGACACCACTGCGATAGTGGAGCCCAAGAAATATACCAGAT 1620
|||||
Db 1561 CAAGCTTGATCATGATGACACCACTGCGATAGTGGAGCCCAAGAAATATACCAGAT 1620
QY 1621 GCTGCTTGATCATGATGATCTGACAGAAATCCATGCTGTGGAGCTCTTATTGTTAT 1680
|||||
Db 1621 GCTGCTTGATCATGATGATCTGACAGAAATCCATGCTGTGGAGCTCTTATTGTTAT 1680
QY 1681 ACGAGGATCCCGGTGTGAGGTGGAGTACTGCAACTGACGCAATGCTCAGACGCAAGAA 1740
|||||
Db 1681 ACGAGGATCCCGGTGTGAGGTGGAGTACTGCAACTGACGCAATGCTCAGACGCAAGAA 1740
QY 1741 GGGAGTGGCGGTGCGGCTCCGACTGTATACCCGGTTCCAAAGCCTAGAGGCTCTTCCGAA 1800
|||||
Db 1741 GGGAGTGGCGGTGCGGCTCCGACTGTATACCCGGTTCCAAAGCCTAGAGGCTCTTCCGAA 1800
QY 1801 CAAGCACCAGTACAGCAAAAGCCTGGGGTGCAGAGTGTCTACATGGTAATGACAGAGT 1860
|||||
Db 1801 CAAGCACCAGTACAGCAAAAGCCTGGGGTGCAGAGTGTCTACATGGTAATGACAGAGT 1860
QY 1861 TATCGAGCACAATACTCCACCACTGTCCAGAGAAACCTGCAAGCTTGCTCATATG 1920
|||||
Db 1861 TATCGAGCACAATACTCCACCACTGTCCAGAGAAACCTGCAAGCTTGCTCATATG 1920
QY 1921 ACACCACTGCGCATAGTGCAGACCCAGAAATACTACCAAAATGCTGGCTTGATATGAAC 1980
|||||
Db 1921 ACACCACTGCGCATAGTGCAGACCCAGAAATACTACCAAAATGCTGGCTTGATATGAAC 1980
QY 1981 TACTGCAAGAAATCCAGATGCTGTGGAGCTCTTATTGTTATGAGAGGATCCCGGTGC 2040
|||||
Db 1981 TACTGCAAGAAATCCAGATGCTGTGGAGCTCTTATTGTTATGAGAGGATCCCGGTGC 2040
QY 2041 AGGTGGAGTACTGCAACCTGACGCAATGCTCAGACGCGAAGAGGAGCTGCCGTGCGCT 2100
|||||
Db 2041 AGGTGGAGTACTGCAACCTGACGCAATGCTCAGACGCGAAGAGGAGCTGCCGTGCGCT 2100
QY 2101 CCGAGTGTACCCGCGTTCCAAAGCCTAGAGGCTCTTCCGAACAAGCACCGACTGAGCAAA 2160
|||||
Db 2101 CCGAGTGTACCCGCGTTCCAAAGCCTAGAGGCTCTTCCGAACAAGCACCGACTGAGCAAA 2160
QY 2161 AAGCCTGGGGTGCAGAGTGTCTACCATGTTAATGAGAGAGTTATCGAGGCATATCTCC 2220
|||||
Db 2161 AAGCCTGGGGTGCAGAGTGTCTACCATGTTAATGAGAGAGTTATCGAGGCATATCTCC 2220
QY 2221 ACGATGTTCACAGGAAGAACTGCCAAGCTTGTCATCTATGACACCACTGCGATAGT 2280
|||||
Db 2221 ACGATGTTCACAGGAAGAACTGCCAAGCTTGTCATCTATGACACCACTGCGATAGT 2280
QY 2281 CGGACCCCGAATACTACCAAAATGCTGGCTTGATGATGAACTCTGACAGAAATCCGAT 2340
|||||
Db 2281 CGGACCCCGAATACTACCAAAATGCTGGCTTGATGATGAACTCTGACAGAAATCCGAT 2340
QY 2341 GCTGTGGAGCTCTTATTGTTATACAGAGGATCCCGGTGTGAGTGGGAGTACTGCAAC 2400
|||||
Db 2341 GCTGTGGAGCTCTTATTGTTATACAGAGGATCCCGGTGTGAGTGGGAGTACTGCAAC 2400
QY 2401 CTGACGCAATGCTCAGACGCGAAGGAGCTCCGTGCGGCTCCGACTGTTTACCCCGGTT 2460
|||||
Db 2401 CTGACGCAATGCTCAGACGCGAAGGAGCTCCGTGCGGCTCCGACTGTTTACCCCGGTT 2460
QY 2461 CCAAGCCTAGAGGCTCTTCCGAACAAGCACCGACTGCGAAGAGGCTGGGGTGCAGAGG 2520
|||||
Db 2461 CCAAGCCTAGAGGCTCTTCCGAACAAGCACCGACTGCGAAGAGGCTGGGGTGCAGAGG 2520
QY 2521 TTGTAACATGGTTAATGACAGAGTTATCGAGGACATATCTCACCACTGTCTACAGGAAGA 2580
|||||

Db	2521	TTGTAACCACTGGTAAATGGACAGAGCTTATCGAGGGACATCTCCACAACCTGTCCACAGAGA	2580
Qy	2581	ACCTGCCAAGGTTGGTCAATCTATGACACACACATCTCCATATGTGGAAACCCAGAAATATAC	2640
Db	2581	ACCTGCCAAGGTTGGTCAATCTATGACACACACATCTCCATATGTGGAAACCCAGAAATATAC	2640
Qy	2641	CCAAATGCTGGCTTGATCATGATCAACTACTGACGAATCCAGATGCTGTGGCAGCTCTTAT	2700
Db	2641	CCAAATGCTGGCTTGATCATGATCAACTACTGACGAATCCAGATGCTGTGGCAGCTCTTAT	2700
Qy	2701	TGTTATACGAGGGATCCCGGTGTCAAGTGGAGATCTGCAACTACCAATGCTCAAC	2760
Db	2701	TGTTATACGAGGGATCCCGGTGTCAAGTGGAGATCTGCAACTACCAATGCTCAAC	2760
Qy	2761	GCAGAGGGACTGCGCTGTGCGGCTCCGACTGTTACCCGGTTCCAAAGCTTAGAGCTCT	2820
Db	2761	GCAGAGGGACTGCGCTGTGCGGCTCCGACTGTTACCCGGTTCCAAAGCTTAGAGCTCT	2820
Qy	2821	TCCCAAGAGACCGACTGTAGCAAAAGCCGGGGTTCAGAGTGGTACCAATGCTAATGA	2880
Db	2821	TCCCAAGAGACCGACTGTAGCAAAAGCCGGGGTTCAGAGTGGTACCAATGCTAATGA	2880
Qy	2881	CAGAGTTATCGAGGCACATCTCCACCACTGTCCACAGAAACCTGCCAAAGCTTGTC	2940
Db	2881	CAGAGTTATCGAGGCACATCTCCACCACTGTCCACAGAAACCTGCCAAAGCTTGTC	2940
Qy	2941	TCTATGACACACACTGCAATGTAGTGGAGCCCGAATATATACCAGAAATGCTGGCTGATC	3000
Db	2941	TCTATGACACACACTGCAATGTAGTGGAGCCCGAATATATACCAGAAATGCTGGCTGATC	3000
Qy	3001	ATGAATCTACTGACAGAAATCCAGATGCTGTGGCAGCTCTTATTTGTTATACGAGGATCC	3060
Db	3001	ATGAATCTACTGACAGAAATCCAGATGCTGTGGCAGCTCTTATTTGTTATACGAGGATCC	3060
Qy	3061	GGTGTCAAGGGAGTACTGCAACTGACAGCAATGCTCAGAGCAAGAGGGACTCCGCTC	3120
Db	3061	GGTGTCAAGGGAGTACTGCAACTGACAGCAATGCTCAGAGCAAGAGGGACTCCGCTC	3120
Qy	3121	GGCGCTCCGACTGTTACCCGGTTCGAAGCTCTTCCTCCGAACAAACACCCAGCT	3180
Db	3121	GGCGCTCCGACTGTTACCCGGTTCGAAGCTCTTCCTCCGAACAAACACCCAGCT	3180
Qy	3181	GAGCAAAAGGCTGGGGTGCAGAGGATCTTCACATGGAATGAGACAAAGTTATCGAGCA	3240
Db	3181	GAGCAAAAGGCTGGGGTGCAGAGGATCTTCACATGGAATGAGACAAAGTTATCGAGCA	3240
Qy	3241	TACTCCACACTGTACAGGAAMAACCTGCAGAACTTGGTCACTATGACACACACACG	3300
Db	3241	TACTCCACACTGTACAGGAAMAACCTGCAGAACTTGGTCACTATGACACACACACG	3300
Qy	3301	CATAGTGGGACCCAGAAATCTACCCAAATGCTGGCTTGATCATGACACTGACGAAAT	3360
Db	3301	CATAGTGGGACCCAGAAATCTACCCAAATGCTGGCTTGATCATGACACTGACGAAAT	3360
Qy	3361	CCAAATGCTGGAGCTCTTATTTGTTATGAGAGGAATCCGGTGTAGAGTGGGAGTAC	3420
Db	3361	CCAAATGCTGGAGCTCTTATTTGTTATGAGAGGAATCCGGTGTAGAGTGGGAGTAC	3420
Qy	3421	TGCAACTGACGCAATGCTCAGACGCAAGAAAGGACTCCGTGGCCTCCGACTGTACC	3480
Db	3421	TGCAACTGACGCAATGCTCAGACGCAAGAAAGGACTCCGTGGCCTCCGACTGTACC	3480
Qy	3481	CCGGTTCCAAAGCCTAAGAGCTCTTCCGAACAAGCACGCACTGACAAAGGCTTGGGGTG	3540
Db	3481	CCGGTTCCAAAGCCTAAGAGCTCTTCCGAACAAGCACGCACTGACAAAGGCTTGGGGTG	3540
Qy	3541	CAGAGTGTCAACATGGTATGAGCAGAGTTATACGAGCACTATCCACACACTGTCA	3600
Db	3541	CAGAGTGTCAACATGGTATGAGCAGAGTTATACGAGCACTATCCACACACTGTCA	3600
Qy	3601	GGAGAAACCTGCCAAGCTTGCTATCTATGACACCACTCGCATAGTGGACCCAGAA	3660
Db	3601	GGAGAAACCTGCCAAGCTTGCTATCTATGACACCACTCGCATAGTGGACCCAGAA	3660

Dd	3601	GGAGAAACCTCCCAAGCTTGGTATCTATGACACCACTCCCAATATGCGAACCCACAA	3660
Qy	3661	TACTACCCAAATGCTGGCTTGATCATGAATCTATGCGAAATCCAGATGCTGTGGCACT	3720
Dd	3661	TACTACCCAAATGCTGGCTTGATCATGAATCTATGCGAAATCCAGATGCTGTGGCACT	3720
Qy	3721	CCTTATGTTATATACGAGGATCCCGGTGTACGGTGGAACTACTGCACCTGACGCAATGC	3780
Dd	3721	CCTTATGTTATATACGAGGATCCCGGTGTACGGTGGAACTACTGCACCTGACGCAATGC	3780
Qy	3781	TCAGACGCGAAGAGGATGCGCGTGGCGCTCCGACTGTTACCCCGGTTCCAAAGCCTTAG	3840
Dd	3781	TCAGACGCGAAGAGGATGCGCGTGGCGCTCCGACTGTTACCCCGGTTCCAAAGCCTTAG	3840
Qy	3841	GCTCCTTCCGAACAGCAAGCACTGAGCAAAAGGCTTGGGGTCCACAGATGCTACCATGT	3900
Dd	3841	GCTCCTTCCGAACAGCAAGCACTGAGCAAAAGGCTTGGGGTCCACAGATGCTACCATGT	3900
Qy	3901	AATGACAGAGTTTTCGAGGCACTACTCCACACTGTACAGGAAGAACTGCCAAGCT	3960
Dd	3901	AATGACAGAGTTTTCGAGGCACTACTCCACACTGTACAGGAAGAACTGCCAAGCT	3960
Qy	3961	TGTCATCTATGACACCACTGCAATGTGGAGCCCAAGATACTACCCAAATGCTGGC	4020
Dd	3961	TGTCATCTATGACACCACTGCAATGTGGAGCCCAAGATACTACCCAAATGCTGGC	4020
Qy	4021	TTGATCATGAACTCTGACAGAAATCCAGATGCTGTGGAGCTCTTATTTATATCAGAG	4080
Dd	4021	TTGATCATGAACTCTGACAGAAATCCAGATGCTGTGGAGCTCTTATTTATATCAGAG	4080
Qy	4081	GATCCCGGTGTCAAGTGGAGTACTGCACATGACCAATGCTCAGACGCAAGAGGACT	4140
Dd	4081	GATCCCGGTGTCAAGTGGAGTACTGCACATGACCAATGCTCAGACGCAAGAGGACT	4140
Qy	4141	GCCGTCGGGCTCCGACTGTATACCCGGTTCGAAGCTGTAGAGGCTCTCCGGAACAACA	4200
Dd	4141	GCCGTCGGGCTCCGACTGTATACCCGGTTCGAAGCTGTAGAGGCTCTCTCCGGAACAACA	4200
Qy	4201	CCGACTGAGCAAAAGGCTGGGGTGCAGAGTGTCTACATGTTAAATGACAGAGTTATCGA	4260
Dd	4201	CCGACTGAGCAAAAGGCTGGGGTGCAGAGTGTCTACATGTTAAATGACAGAGTTATCGA	4260
Qy	4261	GGCACTATATCCACACTGTGCACAGGAAGAACTGCAAGCTTGGTCATCTATGACACCA	4320
Dd	4261	GGCACTATATCCACACTGTGCACAGGAAGAACTGCAAGCTTGGTCATCTATGACACCA	4320
Qy	4321	CACTCGATATGTGGAGCCCGCAATACTACCCAAATGCTGGCTTATCATGAACTACTGC	4380
Dd	4321	CACTCGATATGTGGAGCCCGCAATACTACCCAAATGCTGGCTTATCATGAACTACTGC	4380
Qy	4381	AGGAATCCAGATGCTGTGGCAAGCTCTTATTTATACGAGGGATCCCGGTGTCAAGTGG	4440
Dd	4381	AGGAATCCAGATGCTGTGGCAAGCTCTTATTTATACGAGGGATCCCGGTGTCAAGTGG	4440
Qy	4441	GAGTATGCAACCTGAGCAAGTCTCAGACGCGAAGAGGAGTGGCGTGGCGCTCCGACT	4500
Dd	4441	GAGTATGCAACCTGAGCAAGTCTCAGACGCGAAGAGGAGTGGCGTGGCGCTCCGACT	4500
Qy	4501	GTTATCCCGGTTCCAAAGCTTAGAGGCTCTCTCCGAACAAGCACGACGACCAAGGCT	4560
Dd	4501	GTTATCCCGGTTCCAAAGCTTAGAGGCTCTCTCTCCGAACAAGCACGACGACCAAGGCT	4560
Qy	4561	GGGTTGAGAGATGTCTAATCGTATATGACAGAGTTATCGAGGACATCTCCACACT	4620
Dd	4561	GGGTTGAGAGATGTCTAATCGTATATGACAGAGTTATCGAGGACATCTCCACACT	4620
Qy	4621	GTCACAGGAAGAACTGCAAGCTTGGTATCTATGACACCACTCGCAATATGTGGACC	4680
Dd	4621	GTCACAGGAAGAACTGCAAGCTTGGTATCTATGACACCACTCGCAATATGTGGACC	4680
Qy	4681	CCAGAAATCTACCCAAATGCTGGCTTGATCATGAATCTGACAGAAATCCAGATGCTGTG	4740
Dd	4681	CCAGAAATCTACCCAAATGCTGGCTTGATCATGAATCTGACAGAAATCCAGATGCTGTG	4740

[illegible]

Query Match	Similarity	Score	DB	Length
Best Local Similarity	85.4%	Freid. No. 0		
Matches 3027; Conservative	0	Mismatches 492; Indels 24; Gaps 1;		
322	CCCGGTGTCAGTGGAGACTGTCAACCTGACGCAATGCTCTGACGACGAAAGGACTGCC	381		
Db	1	CCCATGTGAGGTGGGAGTACTGCAACCTGCACATCTCTGACGACGAAAGGACTGCC	60	
QY	382	GTGCGCCTCCGACTGTATACCCGGTTCCAMGCTTAGAGGCTCCTCCGAAACAGCACCG	441	
Db	61	GTGCGACCTCCGAAAGTCAACCCGGTTCCAMGCTTAGAGGCTCCTCCGAAACAGCACCG	120	
QY	442	ACTGAGCAAAAGCCCTGGGGTGCAGAGAGTCTACATGGTAATGAGACAGATTATGAGGC	501	
Db	121	ACTGAGCAAAAGCCCTGGGGTGCAGAGAGTCTACACGTTAATGAGACAGATTATGAGGC	180	
QY	502	ACATATCCACACATGTGTACAGAGAAACCTGCCAAGCTTGGTCTATATGACACACACAC	561	
Db	181	ACATATCTTACACATGTGTACAGAGAAACCTGCCAAGCTTGGTCTATATGACACACACAC	240	
QY	562	TTCGATATGTCCGACCCCGAATACCTACCAATGCTGGCTTATATATGATACCTACCTGACG	621	
Db	241	TCTCATGTGCGAACCCCGAATACTACCAATGCTGGCTTATATGATACCTACCTGACG	300	
QY	622	AATCCAGATGTGCTGGAGCTCTTATGTATTAGAGAGGATCCCGGTGTCAGTGGGAG	681	
Db	301	AATCCAGATCTGTGGAGCCCTTATGTATTACCATGATGCCAATGTACAGTGGGAG	360	
QY	682	TACTGCAACTGACCAATGCTCAACGACGAGAGGACTCCCGTCCGCTCCGACTGT	741	
Db	361	TACTGCAACTGACCAATGCTCAACGACGAGAGGATTCGCTCACCTCTGACTGT	420	
QY	742	ACCCGGTTCACAGGCTTAGAGGCTCCTTCCGAAACAGCACCGACTGAGCAAAAGGCTGG	801	
Db	421	ACCCGGTTCACAGGCTTAGAGGCTCCTTCCGAAACAGCACCGACTGAGCAAAAGGCTGG	480	
QY	802	GTGCGAGGTGTACCATGTATATGAGACAGATTATGAGGACATACTCCACACTGTC	861	
Db	481	GTGCGAGGTGTACCATGTATATGAGACAGATTATGAGGACATACTCTACCACTGTG	540	
QY	862	ACAGGAAGAAGCTGCAAGCTTGTGCATCTTAGAACACACATCTGCATATGTGACACCCA	921	
Db	541	ACAGGAAGAAGCTGCAAGCTTGTGCATCTTAGAACACACATCTTGCATATGTGACACCCA	600	
QY	922	GAATACTACCCAAAGTGTGATTCATGAATCTACTCAGCAATCCAGATCTGTGCA	981	
Db	601	GAATACTACCCAAATGTGTGATTCATGAATCTACTCAGCAATCCAGATCTGTGCA	660	
QY	982	GCTCCTTATTTATACGAGAGGATCCCGGTGTCAAGTGGGAGTACTGCAACTGACGCA	1041	
Db	661	GCCCTTATTTATACGATGTATCCCAATGTCAAGTGGGAGTACTGCAACTGACACAA	720	
QY	1042	TGCTCAGACGCGAAGAGGACTCGGTGCGGCTCCGACATGTAACCCGGTTCCAGCCTA	1101	
Db	721	TGCTCAGACGCGAAGAGGACTCGGTGCGGCTCCGACATGTAACCCGGTTCCAGCCTA	780	
QY	1102	GAGGCTCTTCCGAAACAGCACCGACTGACGAAAGGCTGGGGTCCAGAGTGTACCAT	1161	
Db	781	GAGGCTCTTCCGAAACAGCACCGACTGACGAAAGGCTGGGGTCCAGAGTGTACCAT	840	
QY	1162	GGTATATGACAGAGTTATTCGAGGCAATATCTCACCACTGTTCACGAGGAAGAACTGCCAA	1221	
Db	841	GGTATATGACAGAGTTATTCGAGGCAATATCTCACCACTGTTCACGAGGAAGAACTGCCAA	900	

QY	1222	GCATTGTCATCTATGACACACACACTGGCATATGCGACCCCAAGATACTACCCAATATGCT	1281
Db	901	GCTTGGTCATCTATGAAACACACTCTCATATGCGACCCCGGAAACCTACCCAATATGT	960
QY	1282	GGCTTGAATGAAACTACTACAGAAATCCAGATAGCTGTGGACGTCTTATTTATATACG	1341
Db	961	GGCTTATCAAGAACTACTACAGGAATCCAGATCCTGTGGACGCCCTTATTTATATACG	1020
QY	1342	AGGGATCCGGGTGTCAAGTGGAGTACTGCACCTGACGCAATGCTCAGACGACAGAGG	1401
Db	1021	ATGATATCCCATATGTCAAGTGGAGTACTGCACCTGACCAATGCTCAGACGACAGAGG	1080
QY	1402	ACTGCGCTGCGGCTCCGACGTGTACCCGGTTCGAAGCTGTAGAGGCTCTTCGGAACA	1461
Db	1081	ACTGCGCTGCGACCTCCGATGTGTACCCGGTTCGAAGCTGTAGAGGCTCTTCGTAGCAA	1140
QY	1462	GCACCGACTAGCAAAAGCCTGGGTGTGCAAGAGTGTCTACCATGATATGACAGAGTAT	1521
Db	1141	GCACCACTAGCAAAAGCCTTGGGTGTGCAAGAGTGTCTACCAAGTATATGACAGAGTAT	1200
QY	1522	CGAAGCACATATCTCCACCACTGTGCACAGGAAGAACTGGCAAGCTTGGTTCATCTATGCA	1581
Db	1201	CGAAGCACATATCTCACCACTGTGCACAGGAAGAACTGCAAGCTTGGTTCATCTATGACA	1260
QY	1582	CCACATCTCGATATGTGCGAGCCCGAAGTACTACCCAAATGCTGGCTGTATCAATGACATC	1641
Db	1281	CCACATCTCATATGTGCGAGCCCAAGAAATCTACCCAAATGCTGGCTTGGTCAAGAACTAC	1320
QY	1642	TGCAGAAATCCAGATGCTGTGGCACTCCTTATTTATACAGAGGATCCGGGTTCAGG	1701
Db	1321	TGCGGAATCCAGATCTCTGTGGCAAGCCCTTGGTGTATATCAACCGATCCCAAGTGTTCAGG	1380
QY	1702	TGGAGTACTGCAACCTGCAGCAATGCTGACAGCGCAAGAGGAACTGCGTGCAGCTCGG	1761
Db	1381	TGGAGTACTGTCAACCTGCACACATGCTGACAGTGCAGAGGAACTGTCTATGCTCTCA	1440
QY	1762	ACTGTTACCCGGTTCCAAGCCTAGAGGCTCCTTCGGAAGAAGCACCGACTGACGAAGG	1821
Db	1441	AATTTTATTCGGTTCCAACCTTAGAGGCTTTCTTGCAACAAGAACTACTAGAGAAAC	1500
QY	1822	CCGTGGGTGCAAGAGTGTCTACCATGATATGACAGAGTATGAGGCAACATCTCCAC	1881
Db	1501	CCCGGGTACAGAGTGTCTACTACATATATGACAGAGTATGAGGCAACATCTCCAC	1560
QY	1882	ACTGTCAAGAAAGAACCTCCCAAGCTTGGTCACTATATACACCAACACTCGCATATGTGG	1941
Db	1561	ACTGTTCAGGAAGAACTTCCCAAGCTTGGTCACTATATACACCAACAGCATATGTGG	1620
QY	1942	ACCCCAATATATACCAATGCTGGCTGTGATCATTAATCTAGTCAGAGAAATCCAAATGCT	2001
Db	1621	ACCCCAAAAATCTATCCAAATGTGTGCTGTACCAAGAACTACTGTGAGGAATCCAAATGCT	1680
QY	2002	GTGGCAGCTCTCTATTGTTTATACAGAGGATCCCGGTGTGAGTGGGAGTACTGCAACTG	2061
Db	1681	GAGATTCGCCCTGTGTGTATATACATGATCCCAAGTGTCAAGGTGGGAGTACTGCAACTG	1740
QY	2062	ACGCAATGTCTACAGCGAAGAGGACTGCCGTGGCCCTCCGACATGTTATACCCGGTTCGA	2121
Db	1741	ACACAAATGTCTGTGACAGAAATCAAGTGTCTTGAAGATCTTACAAATGTGTCCCAATCTCA	1800
QY	2122	AGCCTTAGAGCTCTTCTCGAAGAACGACGACTGACCAAGGCTCTGGGTGTGAGAGTGC	2181
Db	1801	AGCACAAGAGCTTCTTGTGAAGAACCAACAGCGAACAAGTCCGAGGTCCAGAGATGCG	1860
QY	2182	TACCATGGTAAATGACAGAGTTATCGAGGCATACTCCACACTGTCTCAGAGAAAGAC	2241
Db	1861	TACCATGGTGTATGACAGAGTTATCGAGGCTCAATCTCCACACATGTCTCAGAGAGACA	1920
QY	2242	TGCCAATGCTGTATATATAGACACCAACTGSCATATGCTGAGCCGACCCGAATATCATACCA	2301
Db	1921	TGTCAATGCTGTGTCTCTATAGAACCAACTGTGCATACAGAGACAAACAATATATATCA	1980
QY	2302	AATCTGTGGCTTGAATCAGAACTACTGACAGAAATCCAGATGCTGTGGACAGCTCTTATTTGT	2361

Dd	1981	GATGCTGGCTGACCGAGGAACCTACTGCSAGAACTCCAGATGCTGAGATTCGGCCCTGGGT	2040
Qy	2362	TATAGAGAGATCCCGGTGTCAAGTGGGAGTATGCTCAACTGACGCAATGCTCAGACGCA	2421
Dd	2041	TATACATGATGCCACGTGTCAAGTGGAGTATGCAACCTGACACATGCTCCAGTGCA	2100
Qy	2422	CAAGGAGCTGCGCTGCGGCTCCGACTGTTACCCCGTTCCAAAGCTAGAGGCTCTCTCC	2481
Dd	2101	GAATCAATGTCTCTCGCAAGTCA-----TGGCTGTTCT	2136
Qy	2482	GAACAGACCGACTGAGCAAAAGCGTGGGGTGCAGAGATGCTTACAGTGTAATGACAG	2541
Dd	2137	GAACAGACCAATGAGCAAAAGCCCGGGGTCCAGAGCTGCTTACCATGTGATGAGACAG	2196
Qy	2542	AGTATGAGGCAATATCTCCACCCTGTGCACAGAAAGAACTGCCAAGCTTGGTCACT	2601
Dd	2197	AGTTATCGAGGTATCTCTCCACCCTGTGCACAGAAAGATATGCACTTGGTCTCT	2256
Qy	2602	ATGACACCACTGCACTAGTCGAGCCCCGAAATATCTCCAAATGCTGGCTGATCATG	2661
Dd	2257	ATGACACCACTGCACTAGTCGAGACCATTAGAAATCTACCCAAATGCTGGCTGACCAAG	2316
Qy	2662	AACTACTGCAAGAACTCCAGATGCTGTGGCAGCTCTTATTTATGACAGGAGATCCGGT	2721
Dd	2317	AACTACTGCAAGAACTCCAGATGCTGTGGAGATTCGCTTGTATATCCATGATCCGAGA	2376
Qy	2722	GTCAAGTGGAGTCACTCAACTGACGCAATGCTGACGCAAGAAAGGACTGGCGTCCG	2781
Dd	2377	GTCAATGAGGAGTCACTCAACTGACCAATGTGTGATGAAATCAAATGTCTCTTCA	2436
Qy	2782	CTCTGACTGTTAACCCCGGTTCCAAAGCGTATGAGGCTCTTCCGAAACAGCACCACTGAG	2841
Dd	2437	ACTCCATGHTGHTCCAGATTCCAAGAGAGAGTTCCTTGAAAGAGCACCAACTGAA	2496
Qy	2842	CAAAAGCGTGGGGTGCAGAGTGTCACTCAATGTGTAATGACAAAGTATTCAGAGCAATAC	2901
Dd	2497	AAACACCTCGGGGTGCAGAGCTGTCAACAAAGTGATGACAAAGTATTCAGAGCAATTC	2556
Qy	2902	TCACACACTGTCAAGAGAAAGCAACTGCAAGCTTGGTCACTATGACACCACTCGCAT	2961
Dd	2557	TCACACACTATTCAGAGAAAGCAATGTCACTCTGGTGTCTATGACACCACTCGCAT	2616
Qy	2962	AGTGGAGCCCAAGAAATCTACCAATGCTGTGCTATATTACTACTGCAAGAAATCA	3021
Dd	2617	CGGAGATCCCATTAAGCTATCCAAATGCTGGCTCACAGGAACATATGCAAGAAATCA	2676
Qy	3022	GATGCTGTGGAGCTCCCTTATTTATTTATGACAGGAGTCCCGGTGAGTGGAGATGAC	3081
Dd	2677	GATGCTGTGGATTCGCTCTGTGTGTTACCATGTGATCCCAAGTGTCAAGTGGGAGTACGC	2736
Qy	3082	AACTGACGCAATGCTCAAGACGCAAGAGGACTGCGTGGCGCTCCGACTGTTACCCG	3141
Dd	2737	AACTGACGCAATGCTCCAGTAGACAGTAAGTGTCTTACAACTCCACAGGTGTCTCCG	2796
Qy	3142	GTTCCAACCTTAGAGGCGCTTCGGAACAAGCAACGACTGACCAAGGCTGGGGTGCAG	3201
Dd	2797	GTTCCAACCAAGAGGCTCTCTTGAAACAAGCAACACTGAGAAAAAGCCCTGTGTTCCAG	2856
Qy	3202	GAGTCTCAATGTTAATGACAGAGTTATGAGGACACTCTCCACACTGTGCACAGA	3261
Dd	2857	GATGCTCAATGTTAATGACAGAGTTATGAGGACACTCTCCACACTGTGCACAGA	2916
Qy	3262	AGAACCTGCAAGCTGTGATATGATGACACCACTGCAATGTGAGAACCCCGAATAC	3321
Dd	2917	AGAACCTGCAAGCTGTGATATGATGACCACTGCAATGTGAGAACCCCGAATAC	2976
Qy	3322	TACCCAAATGCTGCTTATCATTAATCTGACAGAAATCCAGATGCTGTGGCAGCTCT	3381
Dd	2977	TACCCAAATGCTGCTGACAGAAATCTGACAGAAATCCAGATGCTGTGGGAAACACCC	3036
Qy	3382	TATTTGTTATACAGAGATCCCGGTGTGAGTGGAGTACTGCACACTGACGATGCTCA	3441

Db	3037	TGGTGTACAGACTGATCCATGTGTGTAGTGGAGTACTGCAACCTGACACATTCCTCA	3096
QY	3442	GAGCGAAGAGGACTGCCGTGCGCCCTCCAGCTGTTATACCCGGGTCCAAAGCCTAGAGCT	3501
Db	3097	GAACACAGATCATGTGCTTAGAGACTCCACCTGTTGTTCCGGTTCGAAGCATGGAAGCT	3156
QY	3502	CCTTCCGAACAGCAGCAGTGGAGAGGCTGGGGTGTGAGAGATGCTTACCATGTGTAT	3561
Db	3157	CATTGTGAGCAGCAGCAGTGGAGAGGCTGGGGTGTGAGAGATGCTTACCATGTGTAT	3216
QY	3562	GGACAGAGTATTCGAGGCACATCTCCACCCTGTCCACAGAGAATCTGCCAAGCTGG	3621
Db	3217	GGACAGAGTATTCGAGGCACATCTCCACCCTGTCCACAGAGAATCTTCAATCTTGG	3276
QY	3622	TCATCTATGACACACACATCTGCAATGCGAGCCCGCAGAAATCTACCAATGCTGGCTTG	3681
Db	3277	TCATCTATGACACACACATCTGCAATGCGAGCCCGCAGAAATCTTGTGACTTG	3336
QY	3682	ATCATGACTCTCTCAGAGAAATCCAGATGCTGTGGCAGCTCTTATTTGTTATACAGAGAT	3741
Db	3337	ACAATGAACTCTCTCAGAGAAATCCAGATGCTGTGGCAGCCCTTGTGTTTACATGAGAC	3396
QY	3742	CCCGGTGTCAAGTGGAGTACTGCAACCTGACGCAATGCTTCAGACGACGAGGACTGCC	3801
Db	3397	CCCGGCTGTCAAGTGGAGTACTGCAACCTGACGCAATGCTTCAGACGAGAGGACTGTG	3456
QY	3802	GTCGGGCTCTGACTGTTTACCCCGGTTTCAGAGCTTAGAGGCTCTTCCGAAAGACGG	3861
Db	3457	GTCACACCTCCGACTGTGTTATCCCGGTTTCAGAGCTTAGAGGCTCTTCTGTACACAGCATCC	3516
QY	3862	ACT 3864	
Db	3517	TCT 3519	
RESULT 3			
LOCUS	HSU19517	2041 bp	mRNA
DEFINITION	Human (apoaargC) long mRNA, complete cds.	linear	PRI 26-JAN-1996
ACCESSION	U19517		
VERSION	U19517.1	GI:642943	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens.		
	Homo sapiens.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 2041)		
AUTHORS	Byrne,C.D., Schwartz,K. and Lawn,R.M.		
TITLE	Loss of a splice donor site at a 'skipped exon' in a gene homologous to apolipoprotein(a) leads to an mRNA encoding a protein consisting of a single kringle domain		
JOURNAL	Arterioscler. Thromb. Vasc. Biol. 15 (1), 65-70 (1995)		
MEDLINE	95268939		
PUBMED	7749817		
REFERENCE	2. (bases 1 to 2041)		
AUTHORS	Lawn,R.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-JAN-1995) Richard M. Lawn, Stanford University School of Medicine, Stanford University Medical Center, Falk Cardiovascular Research Center CV 267, 300 Pasteur Drive, Stanford, CA 94305-5246, USA		
FEATURES			
source	Location/Qualifiers		
	1..2041		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="6"		
	/map="6q27-ter"		
	/tissue-type="Liver"		
	1..2041		
	/gene="apoargC"		
	118..516		
	/gene="apoargC"		
	/standard_name="apolipoprotein (a) related gene C"		

QY	841	GGCACAATCTCCACCCTGTGTACAGGAAGAACTGCGCAAGTTGGTCACTTATATGACACA	900
Db	753	GGCAATCTACTTCCACCACCGTGTACAGGAAGCACTTGGTCCAGAGTTGGTCACTTATATGACACA	812
QY	901	CACATCGCATGTAGTCCGAGCCCCGAAATATACCAATATGCTGGCTTGATATATACATCTGTC	960
Db	813	CACCAGCACAATAGGACTCTCAGAAAAGTACCCAAATGCTGACTTGATCATATACACTACTGTC	872
QY	961	AGGAATCCAGATGCTGTGGCAGCTCTTATTTGTTATACGAGGATATCCCGTGTGAGGTGG	1020
Db	873	AGGAATCCAGATGCTGTGGGGAAG - GCGTGGGTGTTACACGATGAGATCCCAAGTCAAGATGG	931
QY	1021	GAGTACTGCAACCTGAGCGCAATGCTCAGAGGACGAGAAGGAGTATGGCGTGGCGCTCCGACT	1080
Db	932	GAGTACTGCAACCTGAGCGATGCTCAGACACACAGAGGAGGACTGCGAGTGTGCTGTGACT	991
QY	1081	GTTATCCCGGTTTCCAAAGCCTTAGAGGCTCTCTTCCGAAACAGCACCAGACTGAGCAAAAGGCTT	1140
Db	992	GTTATCCCGGTTTCCAAAGCCTTAGAGGATCTTCCAAACACGAGCACAACAGAGGAGGCTT	1051
QY	1141	GGGGTGCAGAGTCTCTACATGTAATGAGAGATTTATGAGGACATCTACTCCACACT	1200
Db	1052	TGGGAGCAGCAGTGTGTACAGGTAATGAGACAGATTATGAGAGGACACTCTTCCACT	1111
QY	1201	GTCACAGGAGAAACCTGCGCAAGCTTGGTGCATCTATGACACACACTCCGATATGTCGAGCC	1260
Db	1112	GTCACAGGAGAAACCTGCGCAAGCTTGGTGCATCTATGACACACACTCCGATATGTCGAGCC	1163
QY	1261	CCAGATATCTACCCAAATGCTGGCTTGATTCATATGAACTACTCCAGGAATTCAGATGCTGTG	1320
Db	1164	-----TGGCCTGTGACAGGAAC - ACTCCAGGAATTCAGATGCTGTGG	1202
QY	1321	GCACCTCTTATTTGTTATATACAGAGGATCCCGGTGTGACAGTGGAGATCTGTCAACCTGACG	1380
Db	1203	AAACAACTCTGTGTATACAAACTGATCCGTGTGTGAGGTGGAGTACTGTCACTGTACA	1262
QY	1381	CAATGCTACAGCAGCAAGGAGACTGCGCTCCGCTCCGACTGTTATACCCCGGTTCCAAAGC	1440
Db	1263	CAATGCTACAGCAAGGAG - TCAAGTGTCTTAAAGACTCCCACTGTTGTTCCCGTTCCAAAGC	1321
QY	1441	CTAGAGGCTCTCTCCGAAACAACACACGACTGAGCAAAAGGCGTGGGAGTGCAGAGAGCTTAC	1500
Db	1322	ATGAGAGCTCATTTGGAAGCAGCACCACCACTGACCAAAACCCCTGTGTGCTCGGACAGTGTAC	1381
QY	1501	CATGTTAATGAGACAGATTATCGAGGCACATCTCTCACACTGTGTCACAGGAAGAACTGTC	1560
Db	1382	CATGTTAATGAGACAGATTATCAAGGCACTTCTCCACACTGTGTCACAGGAAGCAATGTT	1441
QY	1561	CAAGTTGGTATCATATATACACACACACG - CATATGTGGAGCCCCGAAATCTATCCCAAA	1619
Db	1442	CAATTTTGGTCAATCATATACACACACCGGTGTATAGAGAACCCAGAAAATCTATCCCAAA	1501
QY	1620	TGCTGGCTTGTATCATGAACACTACTGCAAGAAATCCAGATGCTGTGGCAGCTCTTATTTGTTA	1679
Db	1502	TGATGCTCTGCAAAATGAACACTACTGCAAGAAATCCAGATGCTGAGCCCTTGGTGTGTT	1561
QY	1680	TACAGAGGATATCCCGGTGTGACAGTGGGAACTACTGCAACCTGACAGCAATCTCATAGACAGCA	1739
Db	1562	TATCCACAACACCCACAGCTCATAGGTGGAGTACTGTGCAACTGTACGCGATCTTCAGA - CAGA	1619
QY	1740	AGGAGCTCCGCTGCGGCTCCGACAGTGTATACCCCGGTTTCCAAAGCTTAGAGAGGCTCTTCCGA	1799
Db	1620	AGGAGCTGTGGTGTGCTCTCCGACATGTTATCCAGGTTCCAAAGCCTTAGAGAGGCTCTTCCGA	1679
QY	1800	ACAAGCA 1806	
Db	1680	ACAAGTA 1686	

ACCESSION	U19518
VERSION	U19518.1
KEYWORDS	GI:642945
SOURCE	.
ORGANISM	Homo sapiens.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 1712)
JOURNAL	Byrne,C.D., Schwartz,K. and Lawn,R.M.
MEDLINE	Loss of a splice donor site at a 'skipped exon' in a gene
PUBMED	homologous to apolipoprotein(a) leads to an mRNA encoding a protein
REFERENCE	consisting of a single kringle domain
AUTHORS	Arterioscler. Thromb. Vasc. Biol. 15 (1), 65-70 (1995)
JOURNAL	9568939
MEDLINE	7749817
PUBMED	2 (bases 1 to 1712)
REFERENCE	Lawn,R.M.
AUTHORS	Direct Submission
TITLE	Submitted (05-JAN-1995) Richard M. Lawn, Stanford University School
JOURNAL	of Medicine, Stanford University Medical Center, Falk
	Cardiovascular Research Center CV 267, 300 Pasteur Drive, Stanford,
	CA 94305-5246, USA
FEATURES	Location/Qualifiers
source	1..1712
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="6"
	/map="6q27-ter"
	/tissue_type="liver"
	1..1712
	/gene="apoargC"
	118..516
	/gene="apoargC"
	/standard_name="apolipoprotein (a) related gene C"
	/codon_start=1
	/protein_id="AA85693.1"
	/db_xref="GI:642946"
	/translation="MEHEVVLILLLEKSAPEETGCSVOECHSNOSRIGTYTV
	TGRICQAWSSMTPIQHSHSTPERKYPNDSGLISCRNPDCSAGPWCYTTPDPMRYECML
	TRGSDDEGVTFVPLTVPVPSLSEDFIOVA"
	1504
misc_feature	/gene="apoargC"
	/note="site of the alternatively spliced variant, see
	Genbank Accession number U19517"
BASE COUNT	490 a 443 c 419 g 360 t
ORIGIN	
Query Match	16.5%; Score 826; DB 9; Length 1712;
Best Local Similarity	74.2%; Pred. No. 9e-246;
Matches 1216; Conservative	0; Mismatches 230; Indels 192; Gaps 6;
Db	1 CTGGGATTGGGACACACTTTCTTGACACTGCTGGCCATGCCAAAATGAACATAAGGAA 60
Db	73 CTGGGATTGGGACACACTTTCTTGCGGCACGTGGCCAGTCCTCCAAAATGAACATAAGGAA 132
Db	61 GTGGTTCTCTCACTCTTTTATTTCTGGAATCAGCAGCACCTGAGCAAAGCATGTGGTC 120
Db	133 GTGGTTCTCTCACTCTGTATTATTCTGAAGTCAGCACCGACCTGAGAAGGGGCTTCTGTG 192
Db	121 CAGGATTGCTACCAGTGGTGATGAGACAGATTATTCAGAGCAGCTACTCCACACTGTACA 180
Db	193 CAGGATGCTACCAGATGATGAGACAGATTATTCAGAGCAGCTACTCCACACTGTACA 252
Db	181 GGAAGAGCTGCGCAAGCTGGTCATCTAATGACACCAACCAATTAATGAGCCACAGAA 240
Db	253 GGAAAGAACCTGCCAAGCTGGTCATCTAATGAGCCACCAACCAAGTAGAAGCCCAAGAA 312
Db	241 AACTACCCAAGATGCGCTTGATCATGGAAGCTACTCAGAGAAATCCAGATGCTGTGCACT 300
Db	313 AAGTACCCAAGATGCGCTTGATCATGGAAGCTACTCAGAGAAATCCAGATGCTGTGCAAGC 372
Db	301 CCTATTGTTATACGAGGAGTCCGGGTCTCAGGTGGAGTACTGCACCTGACGCATATGC 360

Db 373 CCTTGTTTATACGACGATCCAAATGTCAAGTGGAGTACTGCAACCTGACACGCTGC 432
 QY 361 TAGACGACGAAGGAGTCCCTGCCGCTCGACTGTTACCCCGGTTCCAAAGCTTAGAG 420
 Db 433 TCAGACGATGAAGGAGTGTGTTGCTGCTTCGACTGTTATCCAGTTCCAAAGCTTAGAG 492
 QY 421 GCTCCCTCCGACGACACCGACTAGCAAAAGGCTGGGGTGCAGAGTGTACCATGTT 480
 Db 493 GATTATTATACAAAG----- 508
 QY 481 AATGACAGAGTTATCGAGGACATCTCCACCCTGTCACAGAAAGACCTGCCAAGCT 540
 Db 509 ----- 508
 QY 541 TGGTCACTATGACACCACTGCAATAGTCGACCCCAAGTAATACCAATCTGGC 600
 Db 509 -----TGGC 512
 QY 601 TTGATCATGACTACTGCGAAGATCCAGATGCTGTGGCAGCTCTTATTTGTTATACGAGG 660
 Db 513 TTGATCTCCACTACTGCGAAGATCCGATGTTGGCGGCCCTGTGTTATACACG 572
 QY 661 GATCCCGGTGAGTGGGAGTACTGCAACTGACGCAATGCTCAGACGCGACAGAGGACT 720
 Db 573 GATCCCAAGTCAAGTGGGAGTACTGCAACTGACAGGATGCTCAGACAAAGATAGGGCT 632
 QY 721 GCCGTGCGGCTCCGACTGTTACCCGGTTCCAAAGCTTAGAGGCTCTTCCAAACAGCA 780
 Db 633 GTGGCGGCCCTCTGACTATTATCCGGTTCCAAAGCAGAGGATATCTTCCAAACAGCA 692
 QY 781 CCGACTGACAAAGGCTGGGGGAGTGTCTACCATGATGATTAAGTACAGAGATTATGA 840
 Db 693 CTGATTGATCCAAAGGCTTTCATGATGAGTGTCTACCATGAAATGAGACAGATTATCA 752
 QY 841 GGCACATCTCCACACTGTACAGGAGAAAGCTGCCAAGCTTGTCTATATGACACCA 900
 Db 753 GGCACATCTCCACACCGCTACAGGAGAAAGCTGCCAAGCTTGTCTATATGACACCA 812
 QY 901 CACTCCGATAGTGGGACCCCAAGTAATCTACCCAAATGCTGGTGTATCATGAACTACTGC 960
 Db 813 CACCACACAGTGAAGCTCAGAAAGTACCCAAATGCTGACTGATCATGAACTACTGC 872
 QY 961 AGGAATCCAGATGCTGTGGCAGCTTATTTGTTATGACAGGATCCCGGTGACAGTGG 1020
 Db 873 AGGAATCCAGATCTGTGGAG--GCCTTGTTGTTACACGATGATCCCAAAGTCAGATGG 931
 QY 1021 GAGTACTGCAACTGACGCAATGCTCAGACGCAAGAGGACTGCCGCTCCGACT 1080
 Db 932 GAGTACTGCAACTGACACGATGCTCAGACAGAAAGGACTGCAGTGCCTTGACT 991
 QY 1081 GTTACCCCGGTTCCAAAGCTTAGAGGCTCTTCCGAACAAAGCAGCTGACGCAAAAGCCT 1140
 Db 992 GTTATCCCGGTTCCAAAGCTTAGAGGATCTTCCAAACCAAGCAGCAAGGCCAAGGCT 1051
 QY 1141 GGGGTGACAGTGTCTACATGTTATGACAGAGTTATGAGGCACTACTCCACACT 1200
 Db 1052 TGGGAGACAGAGTGTCTACACCGGTAATGACAGAGTTATGAGGCACTACTCCACACT 1111
 QY 1201 GTACAGAGAAAGCTGCGCAAGCTTGGTCACTATGACACCACTGCGATAGTCGGAAC 1260
 Db 1112 GTACAGAGAAAGCTGCGCAAGCTTGGTCACTATGACCGCATATGACGACACA----- 1163
 QY 1261 CCAGAAATACCAAAATGCTGGCTGATCATGACATGACGAGATCCAGATGCTGTG 1320
 Db 1164 -----TGGCTGACAGAGAAC-ACCTGACAGAAATCCAGATTCTGGG 1202
 QY 1321 GCATCTCTTATTTATACGAGGAGTCCGCTGTCAAGTGGAGTACTGCAACTGACG 1380
 Db 1203 AAACAACCTCGTGTTCACACAGCATCCGTGTGAGTGGAGTACTGCAACTGACG 1262
 QY 1381 CAATGCTCAGACGAGAGGACTGCCGTGCGGCTCCGACTGTACCCCGGTTCCAAGC 1440
 Db 1263 CAATGCTCAGAAACAGATC-AGTGTCTTAGAGACTCCACTGTGTTCCGTTCCAAGC 1321

QY 1441 CTAGAGGCTCCTTCGAAACAAGCACCGACTAGCAAAAGCCTGGGTGACAGAGTGTCTAC 1500
 Db 1322 ATGAGAGCTCTATTCTGAAGCAGCACCAACTGAGCAAAACCCCTGTGGTCCGAGTGTAC 1381
 QY 1501 CATGTAATGACAGAGTTATCGAGGACATCTCCACCACTGTACAGAGAAAGCTGC 1560
 Db 1382 CATGTAATGACAGAGTTATCGAGGACATCTCCACCACTGTACAGAGAAAGCTGT 1441
 QY 1561 CAAGCTTGCTATCTATGACACCACTCG-CATAGTCGGAGACCCCAATATCAATCCAA 1619
 Db 1442 CAATCTGGTCAATCATGACACCAACCGGTCATGAGAGACCCCAAGAAATCAACCA 1501
 QY 1620 TGTGCGCTTGATCATGAA 1637
 Db 1502 TGACTCTATGTTTGGAA 1519
 RESULT 5
 AF029691 1649 bp mRNA linear PRI 28-JAN-1998
 LOCUS AF029691
 DEFINITION Papio hamadryas apolipoprotein a (BABAPOA) mRNA, partial cds.
 ACCESSION AF029691
 VERSION AF029691.1 GI:2815615
 KEYWORDS
 SOURCE
 ORGANISM
 Papio hamadryas.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Papio.
 REFERENCE
 AUTHORS 1 (bases 1 to 1649)
 TITLE Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1649)
 AUTHORS Cox,L.A., Jett,C. and Hixson,J.E.
 TITLE Direct Submission
 JOURNAL Submitted (10-OCT-1997) Genetics, Southwest Foundation for Biomedical Research, P.O. Box 760549, San Antonio, TX 78245-0549, USA
 FEATURES
 source
 location/Qualifiers
 1..1649
 /organism="Papio hamadryas"
 /db_xref="taxon:9557"
 <1..1649
 /gene="BABAPOA"
 <1..1366
 /gene="BABAPOA"
 /codon_start=2
 /product="apolipoprotein a"
 /protein_id="BAB97886.1"
 /db_xref="GI:2815616"
 /translation="STVTYGNCCQMSMTPHOHSRTPKNYPNAGLTIRNYCRNPDAEI
 RPKCYTMDPSVREYCNLTQCLVTESSVLETLIVPDPQSASEEPATQSPSEVDC
 YHGGOSYRSGESTYVGRTOQSWSMTPHOHKTPEHNDGLTMYCNPADYGP
 WCPETMDPSVREYCNLTQCLVTESSVLETLIVPDPQSASEEPATQSPSEVDC
 KCKGRVYVGGCVAAHSHRPMYSLRTFGKHFGCGTLISPEWTLARCLMSPRPS
 KYVLIGHQEVNLSHVOELVSKLSEEPGADIALKLSRPALITDKYLPAQLSPN
 YVITAMTECYLTGMEETQGTGAGLLEFARLPVLENTCNRYELNGRVSTELCAQH
 LAGTDSQGDGSGPVVCFDKDKYILKGIITSMGPGCARPNKPGYVAVSSFTWIRGV
 MRNN"
 2..91
 /gene="BABAPOA"
 /note="exon A; encodes kringle IV 9a"
 92..271
 /gene="BABAPOA"
 /note="exon B; encodes kringle IV 9b"
 272..434
 /gene="BABAPOA"
 /note="exon C; encodes kringle IV 10a"
 435..614
 exon
 exon
 exon


```

|||||
Db 301 CCTATTGTTATGAGTGGATCCCATGTGAGTGGAGTGCACACCTGACATGCG 360
Qy 361 TCAGACACGAGAGGACTGCGCTGCGCGCTCCACTGTTCACCCGGTCCAGACCTAGAG 420
Db 361 TCAGATGACAGAGGAGACTGCGCTGCGCACCTCCGAAATGTCACCTGGTTCACACCTAGAG 420
Qy 421 GCTCCTTCGCAACAA 435
Db 421 GCTCCTTCGCAACAA 435

RESULT 7
AC027197 194321 bp DNA linear HMG 24-AUG-2002
LOCUS Homo sapiens chromosome 6 clone RP11-189014 map 6, LOW-PASS
DEFINITION
AC027197
SEQUENCE SAMPLING.
AC027197 3 GI:9887745
HMG: HMG5_P1ASFD.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
JOURNAL 2 (bases 1 to 194321)
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 6, clone RP11-189014
Unpublished
2 (bases 1 to 194321)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Gadigan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazares,R., Landers,T., Lehoczkj,J.,
Levine,R., Lien,C., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGuirk,A., McKernan,K., McNeethers,R.,
Medlitz,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,T., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teefaye,S., Theodore,J., Tirelli,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Willson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 194321)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
DeRellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heatford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Larocque,K.,
Lamazares,R., Landers,T., Lehoczkj,J., Levine,R., Lien,C., Liu,G.,
Macdonald,P., Margulis,N., McCarthy,M., McEwan,P., McKernan,K.,
McNeethers,R., Medlitz,J., Meneus,L., Mihova,T., Mienga,V.,
Morrow,J., Murphy,T., Naylor,T., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P.,
Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teefaye,S., Theodore,J., Tirelli,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Willson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

```

TITLE
JOURNAL

COMMENT

```

Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 24, 2000 this sequence version replaced g1:9799792.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: I7625
Center clone name: 189_O_14
-----
* NOTE: This record contains 224 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
742 841: gap of 100 bp in length
842 1594: contig of 753 bp in length
1595 1694: gap of 100 bp
1695 2442: contig of 748 bp in length
2443 2542: gap of 100 bp
2543 3274: contig of 732 bp in length
3275 3374: gap of 100 bp
3375 4104: contig of 730 bp in length
4105 4204: gap of 100 bp
4205 4929: contig of 725 bp in length
4930 5029: gap of 100 bp
5030 5737: contig of 708 bp in length
5738 5837: gap of 100 bp
5838 6574: contig of 737 bp in length
6575 6674: gap of 100 bp
6675 7382: contig of 708 bp in length
7383 7482: gap of 100 bp
7483 8227: contig of 745 bp in length
8228 8327: gap of 100 bp
8328 9072: contig of 745 bp in length
9073 9172: gap of 100 bp
9173 9920: contig of 748 bp in length
9921 10020: gap of 100 bp
10021 10740: contig of 720 bp in length
10741 10840: gap of 100 bp
10841 11566: contig of 726 bp in length
11567 11666: gap of 100 bp
11667 12407: contig of 741 bp in length
12408 12507: gap of 100 bp
12508 13354: contig of 747 bp in length
13255 13354: gap of 100 bp
13355 14095: contig of 741 bp in length
14096 14195: gap of 100 bp
14196 14933: contig of 738 bp in length
14934 15033: gap of 100 bp
15034 15755: contig of 722 bp in length
15756 15855: gap of 100 bp
15856 16584: contig of 729 bp in length
16585 16684: gap of 100 bp
16685 17433: contig of 749 bp in length
17434 17533: gap of 100 bp
17534 18286: contig of 753 bp in length
18287 18386: gap of 100 bp
18387 19127: contig of 741 bp in length

```


[illegible]

Query Match	4.8%; Score 242.4; DB 9; Length 2711;	Best Local Similarity 59.0%; Pred. No. 5.1e-64;	Matches 551; Conservative 0; Mismatches 311; Indels 72; Gaps 5;
Query	AGTGTACACCTGGAATGAGCAGAGTTATGAGGAGACATACCTCCACACAGCTGCACAGAA	868	
Db	874 AGTGTGAGAGGAGAACAGGTGAGAAACTATATCTGGGGAGTGTGGCTTTACGTGTCTGGC	933	
Qy	GAACCTGCCAAGCTTGTCATCTTATGACACCACTCCGATATGTGGACCCAGATACT	928	
Db	934 ACACCTGTACAGGCTGTGAGGCGACACACCCCTTCACACATTAAGAGCACACCAAAAACT	933	
Qy	ACCCAAATGCTGGCTTATCATACACTTACGTGAGAGATACAGATGCTGTGGACCTCTT	988	
Db	994 TTCCCTGAAAAAATTGGATGAAATCTACTGCGCATCTGATGAGAAAAAGGCCCAT	1053	
Qy	ATTGTTATACGAGGATCCCGGTGTAGGTGGGAGTACTGCACCTGACGCAATGCTCAG	1048	
Db	1054 GGTGTATACACCAACCAAGCAAGTCCGCTGGAGATGTAATACCGT-----	1104	
Qy	1049 ACGCAGAAAGGAGCTGCGTCCGCTCCACACTTTTACCCCGGTTTCCAGCCTTAGAGCTC	1108	
Db	1105 -----CTGTAGTCTTCCACATTCACAG-----GAAACAT	1137	
Qy	1109 CTTCGGAAACAAAGCAGCAGCTGAGCAAAAGGCGGGGTGACGAGAGCTTACCATGTATAG	1168	
Db	1138 TGTATCCACAGACACCACTGAGCTTACTCTGTGTGTCCAGAGTGTCTACATGTGTATG	1197	
Qy	1169 GACAGATTATCGAGGCACATCTCCACCACTGTACAGAGAAAGAACTGCCAAGCTTGGT	1228	
Db	1198 GCGCAGAGTACCGAGGCGACATCTCTCCACACCAACCAAGAAAGTGCATGCTGTGT	1257	
Qy	1229 CATCTATACACCACTCGATCTAGTCGAGACCCCAAGTAATCTACCAATGCTGTGTGA	1288	
Db	1258 CATGTATGACACCACTGTGGCATGAGAAGGCCCAAGAAAGTTCCAATGCTGTGGCTGA	1317	
Qy	1289 TCATGAACTACTGAGAGATCCAGATGCTGTGGACGCTCTTATGTATACAGGGATC	1348	
Db	1318 CAATGAACCTACTGAGAGATCCAGATGCGATTAAGAGTCCCTGGGTGTTTACCAAGAAC	1377	
Qy	1349 CCGGTGTAGGTTGGAGTACTGTCAACCTGACCAATGCTTCACAGCAGCAAGGAGCTCCG	1408	
Db	1378 CCAGGTGATGAGTGGAGTACTGTCAACCTGAAAAAATGCTCAAGAAACAGAAAGGAGTGTG	1437	
Qy	1409 TCGGCGCCCGCAGTCTTACCCGGTTCCAAAGCCTTGAAGGCTCCTTCCGAACAAACACGA	1468	
Db	1438 CAGCAGCTCGCGCTGTGGCCCACTTCCAGATGACAGAGCTCTCTTCCGA-----	1486	
Qy	1469 CTGACAAAGGCTGGGGGTGACAGAGTGTACCATGTATATGACAGAGTTATGAGGCA	1528	
Db	1487 -----GGAAGACTGTATGTTTGGGAATGGAAAGATATCCGAGGCA	1527	
Qy	1529 CATACTCACCACTGTACAGAGAAAGAACTGCCAAGCTTGGATCATCTTGCACACACACT	1588	
Db	1528 AGAAGGCAACCACTGTTACTGAGACACCATGCCAGGAATGGCTGCCAGAGGCCCA	1587	
Qy	1589 CGCA---TAGTCGAGCCCAAGAACTACTACCAATGCTGGCTTGATCTGAATCTGCA	1645	
Db	1588 GCCACCGCATTTTCACTCCAGAGAACAAATCCAGGCGAGGTCTGGAAAAAAACTATGACC	1647	
Qy	1646 GGAATCCAGA---TGCTGTGGACGCTCTTATGTATACGGGAGATCCCGGTGTCAAGT	1702	
Db	1648 GTAAACCTCGATGATGTAGTGTGTCTCTGTGTGTACACGCAAAATCCAGAAAAATTT	1707	
Qy	1703 GGGAGTACTGCACCTGACCAATGCTCAGAGCG	1736	
Db	1708 TCGACTACTGATGTCCCTCAATGTGGCGGCTC	1741	

```
RESULT 9
HUMPMGM LOCUS 1784 bp mRNA linear PRI 08-JAN-1995
DEFINITION Human plasmidogen mRNA, 3' end.
ACCESSION K02922
VERSION K02922.1 GI:190112
KEYWORDS 41cytochrome; plasmidogen.
SOURCE Human liver, cDNA to mRNA, library of S.L.Moo.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1784)
AUTHORS Malinowski,D.P., Sadler,J.E. and Davle,E.W.
TITLE Characterization of a complementary deoxyribonucleic acid coding
for human and bovine plasmidogen
JOURNAL Biochemistry 23 (18), 4243-4250 (1984)
MEDLINE 85023311
PUBMED 6148961
FEATURES
source
Location/Qualifiers
1..1784
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="6q26-q27"
1..1784
/gene="PLG"
<1..1784
/gene="PLG"
/feature="PLG"
/feature="pmg mRNA: G00-119-498"
<1..1560
/gene="PLG"
/feature="plasmidogen"
/codon_start=1
/protein_id="AAA60124.1"
/db_xref="GI:387031"
/db_xref="GDB:600-119-498"
/translat="SGHTCOHWSAQTPTHTNRPENPCNKLDENYCRNPDGKRAFPC
HTTNSQVSMREYCKIPSCDSSPVSTEQALAPAPETLPVVOVCYHGDQSYRSTSTT
TGKCOVSMSMTPRHOKTPEPNTNAGITMNYCNPADGPMCFETDTSVREYCNL
KRCSTGASVAPPPVLLPPEVPESEDCMGKRGKRAATYTGPCDMAOE
PBRHSIFTEPTEPRAGLEKNCNPDGSDVGGYTNPRKLYDQVOCAPSDC
GKROVEPKKSGRRVVGCVAHPSHPMOWSLRPFGRHPCGGLISPEVATNACLE
KSPRSYKVLIGAHQPNLEPHVOELVSRLEPFRKDIALLKSSPAVITDKYIP
ACLPSPNVADRTCEFTTGGEGTGFAGLKEAOLPYTEKVCNREPLNGRVOS
TELCAHLAGGTDSQSGSGPLVCFEKDKYIIIGVTSWGLCARPNKRGVYVRSF
VTWIEGVARNN"
BASE COUNT 473 a 452 c 457 g 402 t
ORIGIN 2 bp upstream of HpaII site.
Query Match 4.7%; Score 236.6; DB 9; Length 1784;
Best Local Similarity 65.6%; Pred. No. 3.2e-62;
Matches 410; Conservative 0; Mismatches 179; Indels 36; Gaps 3;
QY 92 CAGCAGCCTGAGCAAGCCATGTGTCAGAGATTGCTACCAAGTGTAGGACAGAGTT 151
DB 221 CAGCAGCCTGAGCAAGCCCTGTGTCAGAGACTGCTACCAAGTGTAGGACAGAGCT 280
QY 152 ATGAGGACAGTCTCCACACAGTGCAGAGAAAGCCGCAAGCTTGTCATCTATGA 211
DB 281 ACCGAGCAGATCTCTCCACACAGCAGAGAAAGTGTACAGCTTGTCATCTATGA 340
QY 212 CACCACTCAACATTAATAGACACAGAAATACCAATGCTGCTGATCATGA 271
DB 341 CACCACTCAACAGAGAAAGCCGCAAGAAATACCAATGCTGCTGATCATGA 400
QY 272 ACTGACAGATCCAGATGCTGAGCTCTTATTTATATAGCGAGGATCCGGTGTCA 331
DB 401 ACTGACAGATCCAGATGCTGAGCTCTTATTTATATAGCGAGGATCCGGTGTCA 460
QY 332 GGTGGAGTACTGCAACATGCTGACGCAATGCTGACGCAATGCTGACGCTC 391
DB 461 GGTGGAGTACTGCAACATGCTGACGCAATGCTGACGCAATGCTGACGCTC 520
```

```
QY 392 CGACHTTACCCGGTTTCAAGCCTAGAGGCTCCTCCGACACAGACGAGTAGCAAA 451
DB 521 CGCCTGTTGCTCTCTCCAGATGTAGAGACTCTTCCAGAAA----- 564
QY 452 GGCCTGGGGTGCAGAGTGTCTACCATGTGTAATGACAGAGTTATGACACATCTCA 511
DB 565 -----GACTGTATGTTGGGAATGGAAAGAAATACGAGACAGAGAGCGCA 610
QY 512 CCACHTGCAGAGAAAGCCAGCTGGTGCATCTATGACACACACCTGCATAG-- 569
DB 611 CCACHTGCAGAGAGCCATCCAGAGCTGCTGCCAGAGAGCCCATAGACACACA 670
QY 570 -TCGACCCCGCAATACTACCAATATGCTGCTGATCATGACTGACAGAAATCCAG 628
DB 671 TTTTCACTCCAGAGAAATCCAGAGCGGGCTGTGAAAAAATATAGCCGTACCTG 730
QY 629 A---TGCCTGGGAGCTCCTTATTTATATGAGGAGATCCCGGTGCAGAGTGTCT 685
DB 731 ATGTGATGTAGTGTGCTCCTGCTGCTACACGACAAATCCAAATCTTACGACT 790
QY 686 GCAACCTGACGCAATGCTCAGAGC 710
DB 791 GTGATGCTCCTCAGTGTGCGGCC 815
RESULT 10
LOCUS AX463622 2433 bp DNA linear PAT 15-JUL-2002
DEFINITION Sequence 1 from Patent WO0250290.
ACCESSION AX463622
VERSION AX463622.1 GI:21886382
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nagai,N., Iaroche,Y. and Collen,D.J.
TITLE A yeast expression vector and a method of making a recombinant
protein by expression in a yeast cell
JOURNAL Patent: WO 0250290-A 1 27-JUN-2002;
THROMB X NV (BR)
FEATURES
source
Location/Qualifiers
1..2433
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2433
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD42143.1"
/db_xref="GI:21886382"
/translat="MEHREYVLLLELTKSGQGERLUDDYVNTGASLSFTYKQLGAG
SIECANACBDEDEFTCRAPQYHSEKQCVMAEKRKSSITIRKRVLYLEKVLISE
CTKNGKNYRGTMSTKNGITQKWSSTSPRPSPATHTSBSLEBNCRANDNDPO
GPMCYTTPDEKRDYCDILECEBECMHSGSNYDGKLSKMSGLECAMOSOPHAG
YIPSKFPNNKLNKNCNPDGSDVGGYTNPRKLYDQVOCAPSDC
KGTGENYRGNAVAVTNGHPCQHWSAQTPTHTNRPENPCNKLDENYCRNPDGKRAFPC
HTTNSQVSMREYCKIPSCDSSPVSTEQALAPAPETLPVVOVCYHGDQSYRSTSTT
TGKCOVSMSMTPRHOKTPEPNTNAGITMNYCNPADGPMCFETDTSVREYCNL
LKRCSSTGASVAPPPVLLPPEVPESEDCMGKRGKRAATYTGPCDMAOE
EPBRHSIFTEPTEPRAGLEKNCNPDGSDVGGYTNPRKLYDQVOCAPSDC
CGKROVEPKKSGRRVVGCVAHPSHPMOWSLRPFGRHPCGGLISPEVATNACLE
KSPRSYKVLIGAHQPNLEPHVOELVSRLEPFRKDIALLKSSPAVITDKYIP
ACLPSPNVADRTCEFTTGGEGTGFAGLKEAOLPYTEKVCNREPLNGRVOS
TELCAHLAGGTDSQSGSGPLVCFEKDKYIIIGVTSWGLCARPNKRGVYVRSF
FTWIEGVARNN"
BASE COUNT 678 a 607 c 616 g 532 t
ORIGIN
Query Match 4.7%; Score 236.6; DB 6; Length 2433;
Best Local Similarity 58.6%; Pred. No. 3.2e-62;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;
```

```
QY 809 AGTGTACCATGTAATGACAGAGTTATGAGGACATACCTCCACACTGTGCAGGAA 868
    |||||
Db 821 AGTGTGAAAGGAAAGGAAACATATGCGGGAATGTGGTGTATCGTTTCCGGGC 860
QY 869 GAACCTGCCAAGCTTGGTTCATCTGTACACACACATGCTGATGTGGACCCAGAAATAT 928
    |||||
Db 881 ACACCTGTGACGACTGTGAGTGACAGACCCCTGCACACATTAACAGACACAGAAACT 940
QY 929 ACCCAAAATGCTGGCTTATCATCATGAACTACTGACAGAAATCCAGATGCTGTGGACCTCTT 988
    |||||
Db 941 TCCCTCCAAAAATTTGGATGAAATATCTGCGCATCTGTGACGAAAAAGGCCCCAT 1000
QY 989 ATTGTATACGAGGATCCCGGTGTGAGTGAGTGAGTACCTGCAACCTGAGCAATGCTCAG 1048
    |||||
Db 1001 GGTGCTCATACACCAACCAACCACTGCGGTGGAGTACTGTAAATATACCTT----- 1051
QY 1049 ACGGAAAGGAGTGCCTGCGCTCGACTGTTACCCCGGTTCCAAAGCTTAGAGCTC 1108
    |||||
Db 1052 -----CCTGTGACTCTCCCGCATCTCCAGTATCCAG----GAACAT 1084
QY 1109 CTTCGGAACAGCACCCGATGAGCAAAAGGCTGGGGTGCAGAGTGTACCATGTGTATG 1168
    |||||
Db 1085 TGGTCCCAACGACACCACTGTAGTAACCTGTGTGTCAGAGTGTCTACCATGTGTATG 1144
QY 1169 GACAGAGTTATGAGGACATATCTCCACACTGTGCAGAGAAAGAACTGCCAAGCTTGT 1228
    |||||
Db 1145 GACAGAGTTATGAGGACATATCTCCACACTGTGCAGAGAAAGAAATGTCTGCTTGT 1204
QY 1229 CATCTATGACACCACTGCTGATAGTGGAGCCCAAGAAATCTACCAATATGCTGGCTTGA 1288
    |||||
Db 1205 CATCTATGACACCACTGCTGATAGTGGAGCCCAAGAAATCTACCAATATGCTGGCTTGA 1264
QY 1289 TCATGAATCTATGAGAGAAATCCAGATGCTGTGCAAGCTCTTATTTATACAGGAGATC 1348
    |||||
Db 1265 CATGAATCTATGAGAGAAATCCAGATGCTGTGCAAGCTCTTATTTATACAGGAGATC 1324
QY 1349 CCGGTGTGAGGTGGAGAGTACTGCAACTGACGCAATGCTCAGACGAGAAAGGAGCTCG 1408
    |||||
Db 1325 CCAAGGTGAGGTGGAGAGTACTGCAACTGAGAAATGCTCAGAGAAAGGAGCTGTG 1384
QY 1409 TCGGGCTTCCGACGTTCATCCCGGTTCCAGAGCTAGAGGCTCTCTCCGAACAGACCGA 1468
    |||||
Db 1385 TAGACCTCTCCGCTGTGCTCTGCTCCAGATGTAGAGACTCTCTCCGAAGAA----- 1437
QY 1469 CTGAGCAAAAGGCTGGGGTGCAGAGTGTACATGTAATGACAGAGTTATGAGGCA 1528
    |||||
Db 1438 -----GACTGTATGTTTGGGAATGGGAAAGATACCGAGGCA 1474
QY 1529 CATATCTCACCACATGTCTACAGAGAAAGAACTGCAAGCTTGGTCAATGATGACACCACT 1588
    |||||
Db 1475 AGAGGGCGACACCTGTACTGTGAGGAGCCATGTCAGAGACTGGGCTGCCAGAGGCCATA 1534
QY 1589 CGCATAG---TCGAGCCCAAAATATACCAATGCTGGCTTATGATGAACATACGCA 1645
    |||||
Db 1535 GACACAGATTTTCACTCCACAGACAAATTCACGGGGGCTGTGAAAAAATTAATCTGCC 1594
QY 1646 GGAATCCAGAA---TGCTGTGGACAGCTCTTATTTATACAGAGAGTCCCGGTCTCAGT 1702
    |||||
Db 1595 GTAACCTGATGTGATGTATGATGTGGTCTGCTGTACACAGCAAAATCCAAAGAAATCTT 1654
QY 1703 GGGAGTACTGCAACCTGACGCAATGCTCAGAGC 1736
    |||||
Db 1655 ACGACTACTGTGATGTCTCCTCAGTGTGGGCCCC 1688

RESULT 11
LOCUS AR037325 2497 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 12 from patent US 5801146.
ACCESSION AR037325
VERSION AR037325.1 GI:5955181
KEYWORDS
SOURCE Unknown.
```

```
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2497)
AUTHORS Davidson, D.J.
TITLE Compound and method for inhibiting angiogenesis
JOURNAL Patent: US 5801146-A 12 01-SEP-1998;
FEATURES
    source
        1. 2497
            /organism="Unknown"
BASE COUNT 691 a 624 c 637 g 545 t
ORIGIN
Query Match 4.7%; Score 236.6; DB 6; Length 2497;
Best Local Similarity 58.6%; Pred. No. 3,2e-62;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

QY 809 AGTGTACCATGTAATGACAGAGTTATGAGGACATACCTCCACACTGTGCAGGAA 868
    |||||
Db 821 AGTGTGAAAGGAAAGGAAACATATGCGGGAATGTGGTGTATCGTTTCCGGGC 929
QY 869 GAACCTGCCAAGCTTGGTTCATCTGTACACACATGCTGATGTGGACCCAGAAATAT 928
    |||||
Db 941 TCCCTCCAAAAATTTGGATGAAATATCTGCGCATCTGTGACGAAAAAGGCCCCAT 1000
QY 989 ATTGTATACGAGGATCCCGGTGTGAGTGAGTGAGTACCTGCAACCTGAGCAATGCTCAG 1048
    |||||
Db 1001 GGTGCTCATACACCAACCAACCACTGCGGTGGAGTACTGTAAATATACCTT----- 1051
QY 1049 ACGGAAAGGAGTGCCTGCGCTCGACTGTTACCCCGGTTCCAAAGCTTAGAGCTC 1108
    |||||
Db 1052 -----CCTGTGACTCTCCCGCATCTCCAGTATCCAG----GAACAT 1084
QY 1109 CTTCGGAACAGCACCCGATGAGCAAAAGGCTGGGGTGCAGAGTGTACCATGTGTATG 1168
    |||||
Db 1085 TGGTCCCAACGACACCACTGTAGTAACCTGTGTGTCAGAGTGTCTACCATGTGTATG 1144
QY 1169 GACAGAGTTATGAGGACATATCTCCACACTGTGCAGAGAAAGAACTGCCAAGCTTGT 1228
    |||||
Db 1145 GACAGAGTTATGAGGACATATCTCCACACTGTGCAGAGAAAGAAATGTCTGCTTGT 1204
QY 1229 CATCTATGACACCACTGCTGATAGTGGAGCCCAAGAAATCTACCAATATGCTGGCTTGA 1288
    |||||
Db 1205 CATCTATGACACCACTGCTGATAGTGGAGCCCAAGAAATCTACCAATATGCTGGCTTGA 1264
QY 1289 TCATGAATCTATGAGAGAAATCCAGATGCTGTGCAAGCTCTTATTTATACAGGAGATC 1348
    |||||
Db 1265 CATGAATCTATGAGAGAAATCCAGATGCTGTGCAAGCTCTTATTTATACAGGAGATC 1324
QY 1349 CCGGTGTGAGGTGGAGAGTACTGCAACTGACGCAATGCTCAGACGAGAAAGGAGCTCG 1408
    |||||
Db 1325 CCAAGGTGAGGTGGAGAGTACTGCAACTGAGAAATGCTCAGAGAAAGGAGCTGTG 1384
QY 1409 TCGGGCTTCCGACGTTCATCCCGGTTCCAGAGCTAGAGGCTCTCTCCGAACAGACCGA 1468
    |||||
Db 1385 TAGACCTCTCCGCTGTGCTCTGCTCCAGATGTAGAGACTCTCTCCGAAGAA----- 1437
QY 1469 CTGAGCAAAAGGCTGGGGTGCAGAGTGTACATGTAATGACAGAGTTATGAGGCA 1528
    |||||
Db 1438 -----GACTGTATGTTTGGGAATGGGAAAGATACCGAGGCA 1474
QY 1529 CATATCTCACCACATGTCTACAGAGAAAGAACTGCAAGCTTGGTCAATGATGACACCACT 1588
    |||||
Db 1475 AGAGGGCGACACCTGTACTGTGAGGAGCCATGTCAGAGACTGGGCTGCCAGAGGCCATA 1534
QY 1589 CGCATAG---TCGAGCCCAAAATATACCAATGCTGGCTTATGATGAACATACGCA 1645
    |||||
Db 1535 GACACAGATTTTCACTCCACAGACAAATTCACGGGGGCTGTGAAAAAATTAATCTGCC 1594
QY 1646 GGAATCCAGAA---TGCTGTGGACAGCTCTTATTTATACAGAGAGTCCCGGTCTCAGT 1702
    |||||
Db 1595 GTAACCTGATGTGATGTATGATGTGGTCTGCTGTACACAGCAAAATCCAAAGAAATCTT 1654
QY 1703 GGGAGTACTGCAACCTGACGCAATGCTCAGAGC 1736
    |||||
Db 1655 ACGACTACTGTGATGTCTCCTCAGTGTGGGCCCC 1688
```


GenCore version 5.1.4-p5-A578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 00:20:02 ; Search time 888 Seconds
(without alignments)
12680.171 Million cell updates/sec

Title: US-09-923-515-3_COPY_1_5000

Perfect score: 5000
Sequence: 1 ctaggacttgagacacattt.....caagcttgcatcatatgac 5000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq-10102:*
1: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	489.6	9.8	924	AA03255	Human angiotensin
2	236.6	4.7	1719	AA86804	Human plasminogen
3	236.6	4.7	1719	AA83545	Nucleic acid encod
4	236.6	4.7	2178	AA580436	DNA encoding novel
5	236.6	4.7	2433	AA77711	Human plasminogen
6	236.6	4.7	2433	ABN89459	Human plasminogen
7	236.6	4.7	2497	AA52284	DNA encoding human
8	236.6	4.7	2497	ABN85296	Human plasminogen
9	236.6	4.7	2679	AA040319	Sequence of a plas

10	236.6	4.7	2732	20	AA35376	SEQ ID 51 of W0991
11	236.6	4.7	2732	24	ABN8196	Human plasminogen
12	236.6	4.7	2753	12	AA012547	Encodes Plasminoge
13	236.6	4.7	2756	12	AA012542	Encodes Plasminoge
14	236.6	4.7	2756	12	AA012548	Encodes Plasminoge
15	236.6	4.7	2756	12	AA012552	Encodes Plasminoge
16	236.6	4.7	2756	12	AA012553	Encodes Plasminoge
17	236.6	4.7	2756	12	AA012554	Encodes Plasminoge
18	236.6	4.7	2756	23	AA566112	DNA encoding novel
19	236.6	4.7	2756	23	AA580439	DNA encoding novel
20	236.6	4.7	2759	12	AA012549	Encodes Plasminoge
21	236.6	4.7	2759	12	AA012550	Encodes Plasminoge
22	236.6	4.7	2759	12	AA012557	Encodes Plasminoge
23	236.6	4.7	2759	12	AA012558	Encodes Plasminoge
24	236.6	4.7	2759	12	AA012543	Encodes Plasminoge
25	236.6	4.7	2762	12	AA012544	Encodes Plasminoge
26	236.6	4.7	2762	12	AA012546	Encodes Plasminoge
27	236.6	4.7	2762	12	AA012545	Encodes Plasminoge
28	236.6	4.7	2765	12	AA012556	Encodes Plasminoge
29	236.6	4.7	2771	12	AA012551	Encodes Plasminoge
30	236.6	4.7	2771	12	AA012555	Encodes Plasminoge
31	236.6	4.7	6020	11	AA006648	Sequence of a DNA
32	235	4.7	2296	14	AA040258	Plasminogen DNA an
33	235	4.7	2296	21	AA089829	Plasminogen DNA an
34	235	4.7	2296	22	AA512747	Human plasminogen
35	235	4.7	2497	12	AB154038	Human plasminogen
36	235	4.7	6010	12	AA011998	Human plasminogen
37	234.6	4.7	2679	18	AA789686	Plasminogen encodi
38	216.6	4.3	1153	24	ABL58257	Canine pro-angiot
39	216.2	4.3	1134	21	AA068202	Human angiotensin
40	216.2	4.3	1154	20	AA068805	Angiotensin KI-4E
41	216.2	4.3	1154	20	AA068805	Angiotensin KI-4E
42	216.2	4.3	1154	20	AA068805	Angiotensin KI-4E
43	216.2	4.3	1154	20	AA068805	Nucleic acid encod
44	206	4.1	2696	20	AA077712	Murine plasminogen
45	202	4.0	273	22	AA03256	Human apoliprote

ALIGNMENTS

RESULT 1
AAD03255
ID AAD03255 standard; cDNA: 924 BP.
AC AAD03255;
XX 13-JUN-2001 (first entry)
XX Human angiotensin inhibitor, LK68 cDNA.
DE Human angiotensin inhibitor; LK68; apolipoprotein(a) kringle domain;
KW angiotensin-mediated disease; cancer; rheumatoid arthritis; therapy;
KW cystostatic; antineumatic; antiarthritic; antipsoriatic; psoriasis;
KW cell migration; ss.
XX Homo sapiens.
OS
XX
FH Key
FT 1.924
FT CDS
FT /tag= a
FT /product= "Human LK68 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
PN W0200119868-A1.
PN 22-MAR-2001.
PD 15-SEP-1999;
PF 15-SEP-1999;
XX 99WO-KR00554.
PR 15-SEP-1999;
PR 99WO-KR00554.

```
XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.
PA
PI Chang J, Kim JS, Park EJ, Yun J, Chung S;
XX
XX WPI: 2001-244787/25.
DR P-PSDB: AAY72944.
XX
XX Novel angiogenesis inhibitor, LK68 for treating angiogenesis-mediated
PT diseases, e.g. cancer and rheumatoid arthritis, has human
PT apolipoprotein(a) kringle domains IV36, IV37 and V38 amino acid
PT sequence -
XX
XX Claim 8; Page 41-42; 50pp; English.
PS
XX The present sequence is a cDNA encoding human angiogenesis inhibitor,
CC LK68 protein. LK68 protein contains the amino acid sequences of human
CC apolipoprotein(a) kringle domains IV36 (LK6 protein), IV37 (LK7 protein)
CC and V38 (LK8 protein). LK68, LK6, LK7 and LK8 are inhibitors are of
CC endothelial cell proliferation, cell migration and normal development of
CC capillaries in the chick embryo chorioallantoic membrane (CAM). LK68
CC protein, its single kringles or their functional equivalents, are useful
CC for treating angiogenesis-mediated diseases, such as cancer, rheumatoid
CC arthritis, psoriasis or ocular angiogenic disease in animals or humans.
CC LK68 is useful as an anticancer agent and also for inhibiting primary
CC tumour growth.
XX
XX Sequence 924 BP; 256 A; 252 C; 225 G; 191 T; 0 other:
SQ
Query Match 9.8%; Score 489.6; DB 22; Length 924;
Best Local Similarity 83.0%; Pred. No. 3e-142;
Matches 558; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
OY 107 AAAGCATGTCGTCAGAGATTTGCTACCATGTGTGAGACAGACTTATCCAGGACACT 166
DB 2 AAAGCCCTGTGTCAGAGATTTGCTACCATGTGTGAGACAGACTTATCCAGGACACT 61
OY 167 CCACACTGTTCACAGAGAGACCTGCCACCTGTGTCATGTATGACACCATCAACATA 226
DB 62 CCACACTGTTCACAGAGAGACCTGCCACCTGTGTCATGTATGACACCATCAACATA 121
OY 227 ATAGACACACAGAAACTACCCAAATGCTGCTTATGATCAACTACGAGGAATCCAG 286
DB 122 AGAGAGACCCAGAAACTACCCAAATGCTGCTTATGATCAACTACGAGGAATCCAG 181
OY 287 ATGCTGTGACAGCTCTTATGTATATACAGAGGATCCCGCTGTCAGGTGGAGATCTGCA 346
DB 182 ATTTGTGGAAACACACCTGTGTATACACACCATCCGTGTGTGAGGTGGAGATCTGCA 241
OY 347 ACCTGACGCAATGCTGACAGCAGAAAGGACTCCGTCGCGCTCCGACTGTATCCCGG 406
DB 242 ATCTGACACATGCTCAGAAACAGAAATCAAGTGTCTCTAGAGACTCCACACTGTGTCCAG 301
OY 407 TTCCAACTCAGAGGCTCTTCGGAACACGACGACGATGAGCAAGGCTGGGGGTCCAG 466
DB 302 TTCCAACTCAGAGGCTCTTCGGAACACGACGACGATGAGCAAGGCTGGGGGTCCAG 361
OY 467 AGTGTACCATGTGTATGACAGAGTATGAGGACACATCTCCACACTGTCTACAGAA 526
DB 362 AGTGTACCATGTGTATGACAGAGTATGAGGACACATCTCCACACTGTCTACAGAA 421
OY 527 GAACCTGCCAAGCTTGCTCATGTATGACACACATCTCCATAGTCGGACCCCGAATACT 586
DB 422 GGAATCTCAATCTTGCTCATGTATGACACACATCTCCATAGTCGGACCCCGAATACT 481
OY 587 ACCCAATGCTGCTGTGATCATGAACTACTGACAGAAATCCAGATGCTGTGAGCTCTT 646
DB 482 ACCCAATGATGCTGTGATCATGAACTACTGACAGAAATCCAGATGCTGTGAGCTCTT 541
OY 647 ATTGTATTAGAGGAGATCCCGGTGTGAGGTGGAGTACTGCAACTGACGCAATGCTCAG 706
DB 542 GGTGTATTACACGAGCCCGACATCAGGTGGAGTACTGCAACTGACGCGATGCTCAG 601
```

```
OY 707 ACGCAGAAAGGACGTGCTCGGCTCCGACTGTATACCCCGTTCCAGACCTAGAGCTC 766
DB 602 ACACAGAAAGGACGTGCTCGGCTCCGACTGTATACCCCGTTCCAGACCTAGAGCTC 661
OY 767 CTTCCGAAACAG 778
DB 662 CTTCTGAAACAG 673
RESULT 2
AXX86804
ID AAX86804 standard; cDNA; 1719 BP.
XX
XX AAX86804;
AC
XX 20-SEP-1999 (first entry)
DE
XX Human plasminogen cDNA.
KM Angiostatin; growth inhibition; endothelial cell; angiogenesis;
KW tumour; solubilisation; bacterial inclusion body; plasminogen; ss.
OS Homo sapiens.
PN M0935248-A2.
PD 15-JUL-1999.
PF 11-JAN-1999; 99WO-US00048.
PR 12-JAN-1998; 98US-0071247.
XX
XX (SEAR ) SEARLE & CO G D.
PA Casperson GF, Polazzi JO, Violand BN;
PI WPI: 1999-430389/36.
DR
XX Producing biologically active angiostatin by solubilizing and
PT refolding recombinant protein, used as anticancer and
PT anti-angiogenic agent
PS
XX Example 1; Page 47-48; 64pp; English.
CC The invention describes the production of angiostatin that comprises (a)
CC culturing cells expressing a gene encoding angiostatin; (b) recovering
CC the gene product; (c) solubilizing and then refolding it at high pH, and
CC (d) isolating the properly folded angiostatin. Angiostatin inhibits the
CC growth and migration of endothelial cells, and their ability to form
CC tubular structures. Angiostatin is used to inhibit growth of tumours
CC and angiogenesis. This method provides efficient solubilisation of
CC angiotensin-containing bacterial inclusion bodies and subsequent
CC refolding to biologically active protein. Sequences AAX86805-824
CC represent specific nucleic acid sequences encoding angiostatin that can
CC be used in the method of the invention. The present sequence represents a
CC human plasminogen cDNA which acts as a template for PCR reactions to
CC amplify angiostatin fragments.
XX
XX Sequence 1719 BP; 507 A; 446 C; 417 G; 348 T; 1 other:
SQ
Query Match 4.7%; Score 236.6; DB 20; Length 1719;
Best Local Similarity 58.6%; Pred. No. 3e-63;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;
OY 809 AGTGTACCATGTGTATGACAGAGTATGAGGACACATCTCCACCATGTCTACAGAA 868
DB 803 AGTGTCTAAGAGGAACAGGTGAAATATATGCGCGGATGTGGCTGTATCCGCGGC 892
OY 869 GAACCTGCCAAGCTTGCTCATGTATGACACACACTGCGCTAGTCGGACCCCGAATACT 928
DB 893 ACACCTGTACACACTGTGAGTGCACAGACCCCTCACAACATTAACAGACACCGAAACT 952
OY 929 ACCCAATGCTGCTGTATCATGAACTACTGACAGAAATCCAGATGCTGTGAGCTCTT 988
```

Db	953	TTCCCTCGCAAAATTTGGATGAAAACTACTGCGCCGAATCTTGACGGAAAAAGGGCCCCAT	1012
Qy	989	ATTGTTATATACAGAGGATATCCGGGTGTCAGGTGGGAGTACTGCAACCTGACGCAATGCTCAG	1048
Db	1013	GGTGCCATACAAACCAACAGCCCAAGTGGCGGTGGAGTACTGTAAATACCGT-----	1063
Qy	1049	ACGCAGAAAGGAGCTGCCGTGGCCCTCCGACTGTTCACCTCCGGTCCAAAGCTTAAGGCTC	1108
Db	1064	-----CCTGTAGACTCTCCCACTATTCAGG-----GAACAT	1096
Qy	1109	CTTCCGACACAGACCGACTGTAGCAAAAGGCCCTGGGGGTCCAGAGTGGTACATCTGTAATG	1168
Db	1097	TGGCTCCCAAGACCACTGTAGCTTACCCCTGTGTGTCCAGAGCTGCTGCTCATTGTATG	1156
Qy	1169	GACAGAGTTATCGAGGACATATCTACCCACCTGTCACACAGTGTACAGGAAGAACCTGCCAAGCTTGGT	1228
Db	1157	GACAGAGCTTACCGAGGACATCTCTCACACACACAGGAAGAAGTGTAGCTTGGT	1216
Qy	1229	CATCTTATGACACACACCTGCGCTGCGATGCGGAGCCCGAAGTACTACCCAAATCTGTGGCTGA	1288
Db	1217	CATCTTATGACACACACCGGACACAGAGACCCCGAAGAACTACCCCAATCTGTGGCTGA	1276
Qy	1289	TCATGAATCTACTGCAGGAATTCAGATGCTGTGGACGCTCTTATTTGTTATACGAGGATC	1348
Db	1277	CAATGAATCTACTGCAGGAATTCAGATGCGAATAAAGGCCCTGTGTGTTTACCCACAGACC	1336
Qy	1349	CCGGTGTACAGTGGGAGTACTGCAACCTGACCAATGCGACAGACACAGAGGACCTGGC	1408
Db	1337	CCAGCGTCAGAGTGGGAGTACTGCAACCTGTAAAAATGCTCAGAGAACAGAACGAGTGTGG	1396
Qy	1409	TCGGCCCTCCGACTGTTCACCCGGGTTCCAGACCTTAGAGGCTCTTCCGACACAGCACCGA	1468
Db	1397	TAGACCTCCGCTGTGTCTGTCTTCAGATGTAGAGACTCTCTTCCAGAGA-----	1449
Qy	1469	CTGAGCAAAAGGCTGGGGGTGACAGAGTGTCTACATGTTATGACAGAGTTATGAGGCA	1528
Db	1450	-----GACTGTATTTTGGGAATGGGAAGATACCGAGGCA	1486
Qy	1529	CATATCTCACCACTGTGCACAGAGAAGACCTGCCAAGCTTGGTCATCTGTGACACCACT	1588
Db	1487	AGAGGGGAGACCACTGTTCATCGGAGACCCGATGCCAGAGACAGGGGCTCCACGAGAGCCCCATA	1546
Qy	1589	CGCATTA---TCGAGCCGCCAGATTACTACCCAAATGCTGGCTTGATCATGAACACTCTGCA	1645
Db	1547	GACACACATTTTTCACCTCCAGAGACAAATTCACGCGGGGTGTGAAAAAATTACTGCC	1606
Qy	1646	GGAATCCAGA---TGCTGTGGCAGCTCTTATTTGTTATACAGAGGATCCCGGTGTCAAGT	1702
Db	1607	GTAACCTGTAGTGGATGAGTGGGGTCCCTGTGGTCTACACGCAAAATCCAAAGAAAATTTT	1666
Qy	1703	GGGAGTACTGCACACTGACGCAATGCTGCAGAGCG	1736
Db	1667	ACGACTACTGTGATGTCCCMCACTGTGGCGGCC	1700
RESULT 3			
AAAX35345		AAAX35345 standard; DNA; 1719 BP.	
XX	AAAX35345;		
XX	16-JUL-1999	(first entry)	
DE		Nucleic acid encoding a multifunctional protein of the invention.	
XX			
KW		Angiostatin; endostatin; interferon; thrombospondin;	
KW		interferon-inducible protein; platelet factor 4; anti-angiogenic;	
KW		anti-tumor; multifunctional protein; angiogenic-mediated disease;	
KW		cancer; diabetic retinopathy; macular degeneration; arthritis;	
XX		tumor cell production; ss.	
OS		Synthetic.	

OS	Homo sapiens.
XX	PN WO916889-A1.
XX	PD 08-APR-1999.
XX	PF 30-SEP-1998; 98WO-US20464.
XX	PR 01-OCT-1997; 97US-0060609.
PA	(SEAR) SEARLE & CO G D.
XX	Bolanowski MA, Caparon MH, Caspersen GF, Gregory SA;
PI	Klein BK, McKearn JP;
DR	WPI; 1999-255098/21.
PT	New multifunctional proteins useful for treating angiogenic-mediated diseases
PS	Claim 12; Page 80; 121pp; English.
CC	The specification describes multifunctional proteins which comprise combinations of angiostatin, endothelin, interferon, thrombospondin, interference-inducible protein and platelet factor 4, and have anti-angiogenic and/or anti-tumor activity. The multifunctional protein may exhibit useful properties such as having similar or greater biological activity when compared to a single factor or by having improved half-life or decreased adverse side effects, or a combination of these properties. The proteins can be used for treating an angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular degeneration, or arthritis. They can also be used for inhibiting the production of tumor cells (characteristic of lung, breast, ovarian, prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma, hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor growth. The present sequence encodes a multifunctional protein of the invention.
SO	Sequence 1719 BP; 506 A; 445 C; 418 G; 350 T; 0 other;
Query Match	4.7%; Score 236.6; DB 20; Length 1719;
Best Local Similarity	58.6%; Pred. No. 3e-63;
Matches 547; Conservative	0; Mismatches 315; Indels 72; Gaps 5
OY	AGTGTCTACCATGGTAATGACAGAGTTATTGCAGGACATACTCCACCACTGTTCACAGAA 868
DB	833 AGTGTCGAAGGAGAACAGTGAATACTATCGGGGGAATGGCTGTACCGTGTCCGGC 892
OY	869 GAACCTGCCAAGCTTGGTCACTCTATATACACCACTCGCATATGTCGACCACAATATCT 928
DB	893 ACACCTGTCCAGCACTGGAGTGCACAGACCCCCTCACACATPACAGGACACAGAATAACT 952
OY	929 ACCCAAATGCTGGCTTGCATCATGACTACTGTGCAGAGAAATCCAGATGCTGTGCACTCTT 988
DB	953 TTCCTGCAAAAAATTGGATGAATAACTACTGTCCGCAATCTCGACGAAAAGGGCCCCAT 1012
OY	989 ATTGTATTACGAGGAGATCCCGGTGTAGGTGGGAGTACTGCACACTGACGCAATGCTCAG 1048
DB	1013 GGTCGCAATCAACCAACAGCAAGTCCGGTGGGAGTACTGTAAATACCGT----- 1063
OY	1049 ACGCAGAAGGAGTCCGCTGGCCCTCCGCACTGTTACCCCGGTTTCCAAGCTTAGAGCTTC 1108
DB	1064 -----CGTGTGCCTCTCCAGATATCCAGT---GAAAGT 1096
OY	1109 CTTCGGAACAGCAACCCACTGACCAAAAGCCCTGGGTGTCAGSAGTCTACCATGATGATG 1168
DB	1097 TTGGTCCCAACAGCAACCTGACTTAACCCCTGTGGTCCAGACTGCTACCATGATGATG 1156
OY	1169 GACAGAGTTATTCGAGGACATATTCTCCACCACTGTCCAGAGAAAGAACTGCAACCTTGT 1228
DB	1157 GACAGAGCTTACGAGGACATCTCTCCACCAACCAACAGAGAAAGAAAGTCAAGCTTGT 1216
OY	1229 CATCTATACACCACTCTGCGATATGTCGAGACCCAGAAATATCTACCAAAATCTGTGGCTGA 1288

Db 1217 CATCTATGACACCAACCGGACACCAAGACCCAGAAACCAACCAATCTGCTGCA 1276
QY 1289 TCATGAATCTCTCAGAGATCCAGATGCTGTGGACGCTCTTATTGTTATACGAGGATC 1348
Db 1277 CAATGAATCTCTCAGAGATCCAGATGCTGTGGACGCTCTTATTGTTATACGAGGATC 1336
QY 1349 CCGGTGTGAGTGGAGTACTGCAACCTGACCAATGCTGACAGACGAGAGGACGCTGCG 1408
Db 1337 CCAAGCTGAGTGGAGTACTGCAACCTGACCAATGCTGACAGACGAGAGGATGCTG 1396
QY 1409 TCGGCTCTGAGTGGAGTACTGCAACCTGACCAATGCTGACAGACGAGAGGATGCTG 1468
Db 1397 TAGACCTCTGAGTGGAGTACTGCAACCTGACCAATGCTGACAGACGAGAGGATGCTG 1449
QY 1469 CTGAGCAAAAGGCTGGGAGTGGAGTACTGCAACCTGACCAATGCTGACAGAGGATGCTG 1528
Db 1450 -----GACTGTATGTTGGGATGAGGAAAGGATACCGAGCA 1486
QY 1529 CATATCTCACCACCTGTCACAGAAACCTGCCAAGCTTGGTCTATCTATGACCCACT 1588
Db 1487 AGAGGGGACCACTGTTACTGGAGCCCATGCCAGGACCTGGGCTGCCAGGAGCCCATATA 1546
QY 1589 CGCATAG---TCGACCCCAAGATTAATCAATGCTGGCTGTATCATGACTACTGCA 1645
Db 1547 GACACAGCATTTTCACTCCAGAGACAAATCCACGGGCGGTCTGAAAAAATTTACTGCC 1606
QY 1646 GGAATCCAGA---TCGCTGTGGACCTCTTATTGTTATACGAGGATCCGGTCTCAGGT 1702
Db 1607 GTAACTCTGATGTGATGAGTGGTCCCTGCTGTACACGCAATTCAGAAACTTT 1666
QY 1703 GGGAGTACTGCACTGACGCAATGCTCAGAGC 1736
Db 1667 ACGACTACTGTGATGTCTCTCACTGTGCGGCCCC 1700

RESULT 4
AAS80436/c
ID AAS80436 standard; cDNA: 2178 BP.
XX
AC AAS80436;
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX DNA encoding novel human diagnostic protein #16240.
DE
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PsDB; ABG16249.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
PS Claim 1; SEQ ID No 16240; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2178 BP; 453 A; 532 C; 574 G; 619 T; 0 other:
Query Match 4.7%; Score 236.6; DB 23; Length 2178;
Best Local Similarity 58.6%; Pred. No. 3,4e-63;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;
QY 809 AGTGTACCATGATGATAGACAGAGATGATGAGGACCATCTCCACCTGTACAGGAA 868
Db 1250 AGTGTCTAAGAGAAACAGGAAACATATCGGGAATGGGTGTATCCGTGTCGGGC 1191
QY 869 GAACCTGCCAAGCTTGGTCTATATGACACCACTCGATAGTGGACCCAGAAATCT 928
Db 1190 ACACCTGTACAGCTGTGATGACAGACCCCTTACACATATACAGAGAACCAAACT 1131
QY 929 ACCCAATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 988
Db 1130 TCCCTGCAAAATTTGAGTAAATGATGATGATGATGATGATGATGATGATGATGATG 1071
QY 989 ATTGTATACGAGGATCCCGGTGTGAGTGGAGTATGCAACCTGACGCAATCTCAG 1048
Db 1070 GTTGTATACCAACCAACAGCCAGTGGGTGAGTATGATGATGATGATGATGATGATG 1020
QY 1049 ACGCAGAGGATCGCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 1108
Db 1019 -----CTGTGACTCTCTCCCAATATCCAG-----GAACAAT 987
QY 1109 CTTCGAAACAGCAGCAGTACGCAAAAGCGCTGGGTGAGAGGATGATGATGATGATG 1168
Db 986 TGGCTCCACAGCAGCAGTACGCAAAAGCGCTGGGTGAGAGGATGATGATGATGATGATG 927
QY 1169 GACAGAGTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1228
Db 926 GACAGAGTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 867
QY 1229 CATCTATACACCACTCCGATAGTGGACCCCAATATCTACCAATATGCTGCTTGA 1288
Db 866 CATCTATACACCACTCCGATAGTGGACCCCAATATCTACCAATATGCTGCTTGA 807
QY 1289 TCATGAATCTCTCAGAGATCCAGATGCTGTGGACGCTCTTATTGTTATACGAGGATC 1348
Db 806 CAATGAATCTCTCAGAGATCCAGATGCTGTGGACGCTCTTATTGTTATACGAGGATC 747
QY 1349 CCGGTGTGAGTGGAGTACTGCAACCTGACCAATGCTGACAGACGAGAGGATGCTGCG 1408
Db 746 CCAAGCTGAGTGGAGTACTGCAACCTGACCAATGCTGACAGACGAGAGGATGCTGCG 687
QY 1409 TCGGCTCTGAGTGGAGTACTGCAACCTGACCAATGCTGACAGACGAGAGGATGCTGCG 1468
Db 686 TAGCACTCTGAGTGGAGTACTGCAACCTGACCAATGCTGACAGACGAGAGGATGCTGCG 634

QY 1469 CTGAGCAAAAGCCCTGGGGTGCAGAGTGTCTACCATGTGTAATGACAGAGTTATCGAGCA 1528
Db 633 -----GACTGTATGTTGGGAATGGGAAGATACCGAGGCA 597
QY 1529 CATACTCCACACCTGTGACAGAAAGAACCTGCCAAGTTGTGTATCTATACACACACT 1588
Db 596 AGAGGGGACACCTGTTCCTGGAGCGCATGCCAGGACTGGCTCCAGAGAGCCGCCATA 537
QY 1589 CGCATAG---TCGGACCCCAAGTAATACCAATGCTGGCTTGATCAGTACTGCA 1645
Db 536 GACACAGCATTTTCATCTCCAGAGCAAAATCCAGGCGGGCTGGAAAAAATTACTGCC 477
QY 1646 GGAATCCAGA---TGTGTGGCAGCTCTTATTTATAGAGGATCCCGGTCTAGGT 1702
Db 476 GTAACTCTGATGTGATGTAGTGTGCTCCCTGTGTACTACGACAAATCCAAAGAACTTT 417
QY 1703 GGGAGTACTGCACCTGCAGCAATGCTCAGACGC 1736
Db 416 ACGACTACTGTGATGTCTCCCTCAGTGTGGGCC 383

RESULT 5
AAK77711
ID AAK77711 standard; DNA; 2433 BP.
AC AAK77711;
XX 10-AUG-1999 (first entry)
DE Human plasminogen DNA coding region.
XX
XX Plasminogen; human; angiotatin; endostatin; gene therapy; vector;
KM anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;
KM tumour growth; solid tumour; diabetic retinopathy; retina; ss.
XX
OS Homo sapiens.
XX
XX WO926480-A1.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-US24950.
XX
PR 20-NOV-1997; 97US-0975424.
XX
PA (GENE-) GENETIX PHARM INC.
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Bachelet T, Leboulch P, Pawlik RJ;
XX
DR WPI; 1999-357696/30.
DR P-PSDB; AAY08685.
XX
PT Anti-angiogenic gene therapy vectors
XX
PS Disclosure; Fig 5; 83pp; English.
CC This invention describes a novel viral gene therapy vector comprising a
CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
CC from human or murine angiotatin, human or murine endostatin and
CC angiogenesis-inhibiting fusions and fragments, where the viral vector is
CC sufficiently attenuated for use in human gene therapy. The products of
CC the invention have anti-angiogenic, cytostatic, anti-diabetic and
CC ophthalmological activity. The vector is used in gene therapy for
CC inhibiting tumour growth in humans harbouring a solid tumour. The vector
CC expresses an anti-angiogenic polypeptide. An additional use comprises
CC treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
CC inhibits angiogenesis in the vicinity of the retina. The vector is
CC administered to cells ex vivo and then administered to the patient.
XX
SQ Sequence 2433 BP; 678 A; 607 C; 616 G; 532 T; 0 other;

Query Match 4.7%; Score 236.6; DB 20; Length 2433;

Best Local similarity 58.6%; Pred. No. 3,5e-63;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;
QY 809 AGTGTACCATGTGTAATGACAGAGTTATCGAGGACATATCCACCATGCTCAGAGAA 868
Db 821 AGTGTCTGAAGGAAACAGAGTAAACTATCGGGAAATGTGGCTGTATACGTTTCCGGC 880
QY 869 GAACCTGCCAAGCTTGGTATCTATGACACACACTGCATATGTCGAGACCCAGATATCT 928
Db 881 ACACCTGTGAGCAGCTGGAGTGCACAGCCCTCCACACATACAGACACACCAAAACT 940
QY 929 ACCCAATGCTGGCTGTGATTCATGAACTACTGCAGAAATCCAGATGCTGTGGACGCTCT 988
Db 941 TCCCTCGCAAAATTTGGATGAAACTACTCCGCAATCTGACAGGAAAAAGGCCCAT 1000
QY 989 ATTGTATACGAGGATCCCGGTCTCAGTGGAGTACTGCACACTGACGCAATGCTCAG 1048
Db 1001 GGTGCCATFCAACCAACAGCCAAAGTGGGTGGAGTACTGTAAATATACCTG----- 1051
QY 1049 ACGCAGAAAGGACTGCGCTCGAGCTGATACCCGCGTTCCAAAGCCTAGAGGCTC 1108
Db 1052 -----CCTGTGACTCCTCCCATATTCACG-----GAACAAT 1084
QY 1109 CTTCCGACACACACACGACTGACGAAAGGCTGGGTGTCAGAGTGTATCCATGTAATG 1168
Db 1085 TGGCTCCACACACACACACTGTAGTAACTCCGTGTGTCAGAGTGTATCATCATGTGATG 1144
QY 1169 GACAGATTATGAGGACATATCTCCACACTGTTCACAGAGAAACCTGCCAAGCTTGT 1228
Db 1145 GACAGAGTACCGAGGACATCTCCACACACACACAGGAAAGAGTGTAGCTTGTGT 1204
QY 1229 CATCTATGACACACACACCTGATGTCGACCCCGAATATCTACCAATCTGCTGTGA 1288
Db 1205 CATCTATGACACACACACCGGACACAGAAAGCCCGAATATCTACCAATCTGCTGTGA 1264
QY 1289 TCATGAACTACTGACAGGAATCCAGATGCTGTGGCAGCTCTTATTTATACAGGAGATC 1348
Db 1265 CATGAACTACTGACAGGAATCCAGATGCGATTAAGGCGCCCTGTGTTTACCAACAC 1324
QY 1349 CCGGTGTGAGGGGGAGTACTGCAACTGACGCAATGCTCAGACAGCAAGGAGCTGCG 1408
Db 1325 CCAGCGTCAAGGGGAGTACTGCAACTGCAAAATGCTCAGAGACAGAGCGAGTGTG 1384
QY 1409 TCGCGCTCCGAGCTGTACCCCGGTCCAAAGCTTAGAGCTCTTCCGACACACCGGA 1468
Db 1385 TAGCACCCTCGGCTGTGCTGTCTCCAGATGTAGAGACTCTTCCGAGAA----- 1437
QY 1469 CTGAGCAAAAGCCCTGGGGTGCAGAGTGTCTACATGTAATGACAGAGTTATGAGGCA 1528
Db 1438 -----GACTGTATGTTGGGAATGGGAAGATACCGAGGCA 1474
QY 1529 CATACTCCACACCTGTGACAGAAAGAACCTGCCAAGTTGTATCTATGACACACACT 1588
Db 1475 AGAGGGGACACACTGTTCATGGGAGCGCATGSCAGAGACTGGGCTGCCAGAGGCCATA 1534
QY 1589 CGCATAG---TCGGACCCCAAGTAATACCAATGCTGGCTTGATCATGTAATCTGCA 1645
Db 1535 GACACAGCATTTTCATCTCCAGAGCAAAATCCAGGCGGGCTGGAAAAAATTACTGCCC 1594
QY 1646 GGAATCCAGA---TGTGTGGCAGCTCTTATTTATAGAGGATCCCGGTCTAGGT 1702
Db 1595 GTAACTCTGATGTGATGTAGTGTGCTCCCTGTGTACTACAGCAAAATCCAAAGAACTTT 1654
QY 1703 GGGAGTACTGCACCTGCAGCAATGCTCAGACGC 1736
Db 1655 ACGACTACTGTGATGTCTCCCTCAGTGTGGGCC 1688

RESULT 6
ABN89459
ID ABN89459 standard; cDNA; 2433 BP.
XX
AC ABN89459;

```
XX 02-SEP-2002 (first entry)
XX
XX Human plasminogen encoding cDNA SEQ ID NO:1.
DE
XX Human; plasminogen; microplasminogen; miniplasminogen; yeast; vector;
KW yeast expression vector; cardiant; thrombolytic; cerebroprotective;
KW fibrin proteolysis; kringle domain; thromboembolic disease;
KW focal cerebral ischemic infarction; ischemic stroke;
KW arterial thrombotic disease; peripheral arterial occlusive disease;
KW acute myocardial infarction; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..2433
XX FT /tag- a
XX FT /product= "plasminogen"
XX
XX WO200250290-A1.
XX
XX 27-JUN-2002.
XX
XX 20-DEC-2001; 2001WO-BE00217.
XX
XX 21-DEC-2000; 2000GB-0031196.
XX 09-JUL-2001; 2001GB-0016690.
XX 09-JUL-2001; 2001GB-0016702.
XX
XX (THRO-) THROMB-X NV.
XX
XX Collen DJ, Nagai N, Laroche Y;
XX
XX WPI: 2002-500632/53.
XX
XX P-PSDB: ABB81496.
XX
XX Novel expression vector for expressing mammalian plasminogen
XX derivatives in yeast, has nucleotide sequence coding for catalytic
XX domain of plasminogen and/or coding for kringle domains of plasminogen
XX linked to promoter.
XX
XX Claim 6; Page 44-47; 61pp: English.
XX
XX The present invention describes a yeast expression vector (1) comprising
XX a mammalian nucleotide sequence operably linked to a promoter, where the
XX mammalian nucleotide sequence codes for the catalytic domain of
XX plasminogen and further optionally codes for one or more kringle domains
XX of plasminogen, its mutants or hybrids. (1) has cardiant, thrombolytic
XX and cerebroprotective activities, and can mediate fibrin proteolysis.
XX (1) can be used useful for treating a thromboembolic disease in a mammal.
XX The mammalian protein expressed by (1) is useful for treating focal
XX cerebral ischemic infarction (ischemic stroke) or arterial thrombotic
XX diseases such as peripheral arterial occlusive disease or acute
XX myocardial infarction. The present sequence encodes human plasminogen,
XX which is used in the exemplification of the present invention.
XX
XX Sequence 2433 BP; 678 A; 607 C; 616 G; 532 T; 0 other:
XX
XX
XX Query Match 4.7%; Score 236.6; DB 24; Length 2433;
XX Best Local Similarity 58.6%; Pred. No.3.5e-63;
XX Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;
XX
XX 809 AGTGTACCATGTAATGACAGACAGATTATGAGCAGCATATCTCCACCTGTCACAGAA 868
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 821 AGTGTCTGAAGGAACAGTGAACATATCGCGGGAATGGTGTATACCGTTCCGGGC 880
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 869 GAACCTGCCAAGCTTGTCATCTATGACACACACTCGCATAGTGGACCCGAAATCT 928
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 881 ACACCTGTACAGCACTGTGACAGACGCCCTCACACATATAACAGAGAACCCAGAAACT 940
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 929 ACCCAAAATGCTGCTGATCATGAACTGACAGAAATCCAGATGCTGAGAGCTCCTT 988
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 941 TCCTCTGCAAAATTTGGATGAAACTACTGCGCAATCTCTGACGAGAAAGGCGCCCAT 1000
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 989 ATTGTATACGAGGATCCCGGTGACAGTGGAGTACTGCAACCTGACCAATGCTCAG 1048
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1001 GGTGCTATCAACCAACAGCCAGTGGGTGATGTTACTGTAAATACCTT----- 1051
XX
XX QY 1049 ACGCAGAAAGGACTCGCTGCGCTCCGACTGTACCCCGGTCCAGAGCTTAGAGCTC 1108
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1052 -----CCTGTGACTCCTCCCGAGTATCCACG-----GAACAAAT 1084
XX
XX QY 1109 CTTTCGAAACACACCGACTGAGCAAAAGGCTGGGGTGTGAGGTGTATCATGTATATG 1168
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1085 TGGTCTCCACACACACCTACTGAGCTTACCTCTGTGTGCTCCAGAGCTGTACATGTTGATG 1144
XX
XX QY 1169 GACAGATTATGAGAGCAGCATATCTCCACACTGTACAGAGAAAGACCTGCAAGCTTGT 1228
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1145 GACAGAGTACCGGACATCTCTCCACACACAGAGAAAGAGTGTCACTCTTGTGT 1204
XX
XX QY 1229 CATCTATGACACACACAGTGTGATGTGGACCCGAGAAATCTACCCAAATGTGGCTTGA 1288
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1205 CATCTATGACACACACAGGACCCAGAAAGACCCAGAAATCTACCCAAATGTGGCTTGA 1264
XX
XX QY 1289 TCATGAACTACTGACAGGATCCAGATGCTGTGACAGCTCTTATTTATACGAGGATC 1348
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1265 CAATGAACCTACTGACAGGATCCAGATGCGGATTAAGGCGCTGTGTTTACACACAGAC 1324
XX
XX QY 1349 CCGGTGTGAGTGGAGTGTGCAACCTGACAGCAATGCTCAGACGCAAGGAGACTGCGG 1408
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1325 CCAGCTCAGGTGGAGTGTGCAACCTGCAAAATCTCAGAACAGAGAGTGTG 1384
XX
XX QY 1409 TCGGCTCCGACGCTTACCCCGGTCCAGAGCTTCCGAGAGCTCTTCCGAAACAGACGA 1468
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1385 TAGCACTCTCCGCTGTGTGCTGTGCTGAGATGAGAGACTCTCCGAGAGAA----- 1437
XX
XX QY 1469 CTGAGCAAGGCTGGGGTGGAGGCTCTACATGTAATGAGCAGAGTTATCGAGGA 1528
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1438 -----GACTGTATGTTGGGAATGGGAAGATACCGAGGCA 1474
XX
XX QY 1529 CATATCTCACACCTGTACAGAGAAACCTGCAAGCTGTGATCTATGACACACACT 1588
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1475 AGAGGCGCAGCACTGTACTGTAGGAGCGCCATGCCAGACTGGCGCCAGAGACCCCAAT 1534
XX
XX QY 1589 CGCATGAG--TGGAGCCCAAGATCTTACCAATTCGCTGATCATGCACTACTGCA 1645
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1535 GACACAGCAATTTCTACCTCAGAGCAAAATCCAGCGGGGTCTGGAATAAATTAATCTGCG 1594
XX
XX QY 1646 GGAATCCAGA--TCTGTGCGAGCTCTTATGTTATACGAGGATCCGCTGTCAAGT 1702
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1595 GTAACTGTATGTATGTAGTGTGCTGCTGTCTACAGCAAAATCCAAAGAACTTT 1654
XX
XX QY 1703 GGGAGTACTGCAACCTGAGCGCAATGCTCAGAGCG 1736
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 1655 ACGACTACTGTGATGTCCCTCAGTGTGCGGCC 1688
XX
XX
XX RESULT 7
XX AAA52284
XX ID AAA52284 standard; DNA; 2497 BP.
XX
XX AC AAA52284;
XX
XX DB 18-SEP-2000 (first entry)
XX
XX XX DNA encoding human plasminogen, SEQ ID NO:12.
XX
XX KW Plasminogen; human; kringle 5 domain; endothelial cell proliferation;
XX angiogenesis; antiproliferative; antiarteriosclerotic; cyostatic;
XX antipsoriatic; antiinflammatory; antitumor; antirheumatic; antitumor;
XX antitumor; cancer; tumor; autoimmune disease; ss.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT mat_peptide 107..2479
```

FT /tag= a
FT /product= "Human plasminogen"
FT /transl_except= (pos:2150..2152, aa:Ile)
PM US6057122-A.
XX
XX
PD 02-MAY-2000.
XX
XX
PF 05-MAY-1997; 97US-0851350.
XX
XX
PR 03-MAY-1996; 96US-0643219.
PR 03-APR-1997; 97US-0832087.
XX
XX
PA (ABBO) ABBOTT LAB.
XX
XX
PI Davidson DJ;
XX
XX
DR WPI: 2000-349573/30.
DR P-PSDB; AAB01887.
XX
XX
PT Preparation of Kringle five peptide fragment for treating various
PT disorders such as angiogenic, ocular, skin diseases and cancer,
PT involves mixing mammalian plasminogen and elastase followed by
PT incubation and isolation -
XX
XX
PS Disclosure: Fig 3; 48pp; English.
XX
XX
CC The invention relates to a method of preparing plasminogen kringle 5
CC peptide fragments. The method comprises mixing mammalian plasminogen and
CC elastase in the ratio 1:100-1:300, followed by incubating and isolating
CC the fragment. The kringle 5 peptides are inhibitors of angiogenesis and
CC endothelial cell proliferation and migration. The peptides are useful for
CC treating angiogenic diseases, primary and metastatic solid tumours and
CC carcinomas of various organs such as breast, genital tract, endocrine
CC glands, skin, tumours of the brain and eyes and solid tumours arising
CC from haematopoietic malignancies such as leukemias and lymphomas. They
CC are also used for the prophylaxis of various autoimmune diseases (e.g.,
CC rheumatoid arthritis), ocular diseases, skin diseases (e.g., psoriasis),
CC blood vessel diseases (e.g., haemangiomas, Osler-Weber Syndrome),
CC diseases caused by excessive or abnormal stimulation of endothelial cells
CC (e.g., Crohn's disease, atherosclerosis), diseases which have
CC angiogenesis as a pathologic consequence (e.g., cat scratch disease and
CC ulcers). The peptides are also useful as a birth control agent which
CC inhibits ovulation and establishment of the placenta. The present
CC sequence represents DNA encoding human plasminogen.
XX
XX
SQ Sequence 2497 BP; 691 A; 624 C; 637 G; 545 T; 0 other;

Query Match 4.7%; Score 236.6; DB 21; Length 2497;
Best Local Similarity 58.6%; Pred. No. 3; 6e-63;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

QY 809 AGTGGTACCAGTGTATGATGACAGAGATTATGAGGACATATCTCCACCAGTGCACAGAA 868
DB 870 ACTGTCGTGAAGGAAACAGTGAACAACTATCGGCGAATGCGTGTATACCGTCCGGGC 929
QY 866 GAACCTGCCAAGCTTGGTCACTATGACACCACTCGCATCTGGAGCCCAAGAACTACT 928
DB 930 ACACCTGTACGACTGTCACAGACCCCTCACACATACACAGACAGACGAAACT 989
QY 929 ACCCAAAATGCTGGCTGATCATGAATCTGACAGATCCAGATGCTGGGACGCTCCT 988
DB 990 TCCCTGCAAAAATTTGATGAATAAATCTATGCGCCATCTCTGACGAAAAAGGCCCCAT 1049
QY 989 ATTGTTATACGAGGATCCCGGTCTCAGGTGGAGTACTCAACCTGACGCAATGCTCAG 1048
DB 1050 GGTGCAATACCAACCAACGCAAGTGGGAGTCTGTAGATATACGCT----- 1100
QY 1049 ACGCAGAGGAGTCCGCTCGCACTGTTACCGCGGTTCCAGGCTTACAGAGCTC 1108
DB 1101 -----CCTGTGACTCTCTCCCGAGTATCCAGC-----GAAACAT 1133
QY 1109 CTTCCGAACAGCACCAGTACGCAAAAGCCTGGGGTGACAGAGTGTCTACATGTATATG 1168

DB 1134 TGGCTCCACAGACACCTAGAGTAAACCCCTGTGGTCCAGAGTGTGTACCATGTGATG 1193
QY 1169 GACAGAGTATTTGAGGACATATCTCCACCACTGTGTACAGGAAGAACTGCCAAGTTGGT 1228
DB 1194 GACAGAGCTAACCGAGGACATCTCCACACACACAGGAAAGATGTGAGCTTGGT 1253
QY 1229 CATCTATGACACACACCTGCACTAGTCCGACCCCGAATTAATCAACCAATGTGGCTTGA 1288
DB 1254 CATCTATGACACACACCGGACACCGAGAACCCCGAAGAAATTAATCAATGTGGCTTGA 1313
QY 1289 TCATGAACTACTGACGGAATCCAGATGCTGTGGACGCTCTTATTTATACAGGATC 1348
DB 1314 CAATGAACTACTGACGGAATCCAGATGCGGATTAAGATTAACCAACAGACC 1373
QY 1349 CCGGTGTACAGTGGAGTACTGCACTGACGCAATGCTTCAGACGCAAGAGGACCTCG 1408
DB 1374 CCAGCTGACAGTGGAGTACTGCACTGCACTGCAAAATGCTCAGGAACAGAGGATTTG 1433
QY 1409 TCGCGCTCCGACTGTACCCCGGTTCCAAAGCTTAGAGGCTCTCCGAACAGCACCGA 1468
DB 1434 TAGACCTCCGCTGTGTCTGCTCCAGATGTAGAGACTCTCCGAAAGAA----- 1486
QY 1469 CTGAGCAAGAGCCTGGGGTGCAAGAGTGTACCATGTATAGAGACAGATGTATGAGGCA 1528
DB 1487 -----GACTGTATGTTGGGAATGGAAAGATACCGAGGCA 1523
QY 1529 CATATCCACACTGTACAGAGAAAGACCTGTCATCTATGATACACCACT 1588
DB 1524 AGAGGGGACACACTGTATCTGTGGAGCCATGCGAGACTGGGCTGCCAGAGCCCATTA 1583
QY 1589 CGCATG---TCGACCCCAAGATTAACCCAAATGCTGCTGTATCATGAACTACTGCA 1645
DB 1584 GACACGCACTTTTCACTCCAGAGAAATCCAGGCGGGCTGTGAAAAAATTAATGCGC 1643
QY 1646 GGAATCCAGA---TGCCTGGCAGCTCTTATTTATTAACAGGAGATCCCGTGTACGT 1702
DB 1644 GTAACTGTATGATGTATGAGTGTGCTGTGTACACGAAATCCAAAGAAACTTT 1703
QY 1703 GGGAGTACCTCAACCTGACGCAATGCTCAGACGC 1736
DB 1704 ACGACTACTGTATGTCTCTCTCATGTGCGGCC 1737

RESULT 8
ABN85296
ID ABN85296 standard; DNA; 2497 BP.
XX
XX
AC ABN85296;
XX
XX
DT 30-SEP-2002 (first entry)
XX
XX
DE Human plasminogen coding sequence.
XX
XX
KW Human: antirheumatic; antiarthritis; gene therapy; anti-angiogenic;
KW rheumatoid arthritis; plasminogen; angiotatin; gene; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 50..2482
FT /tag= a
FT /product= "Human plasminogen"
FT sig_peptide 50..103
FT /tag= b
PN WO200253191-A1.
XX
XX
PD 11-JUL-2002.
XX
XX
PE 03-JAN-2002; 2002MO-KR00001.
XX
XX
PR 05-JAN-2001; 2001KR-0000691.

XX (VIRO-) VIROMED LTD.
 PA
 XX Kim J, Ho S, Park E, Kim S;
 PI
 XX WPI: 2002-583596/62.
 DR P-PSDB: ABB83470.
 XX
 XX Novel composition for gene therapy against rheumatoid arthritis,
 PT comprising a DNA encoding anti-angiogenic protein or its parts,
 XX
 XX Disclosure; Page 48-54; 84pp: English.
 ES
 XX The present invention relates to a composition for gene therapy,
 CC comprising a DNA encoding an anti-angiogenic protein, which shows
 CC therapeutic effects on rheumatoid arthritis. The composition is useful
 CC for treating rheumatoid arthritis and the gene therapy is effective,
 CC lasting for 14 days. The present sequence is the coding sequence for
 CC human plasminogen. Angiostatin, which consists of residues 98-440 of
 CC plasminogen, was used as an anti-angiogenic protein.
 XX
 XX Sequence 2497 BP; 691 A; 624 C; 637 G; 545 T; 0 other;

Query Match 4.7%; Score 236.6; DB 24; Length 2497;
 Best Local Similarity 58.6%; Pred. No. 3.7e-63;
 Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

OY 809 AGTGTACCATGTAATGACAGAGTTATCGAGGACACATCTCCACCTGTCCACAGAA 868
 DB 870 AGTGTGAAAGGAAACGTAATCGCGGAATGTGCTGTACCGTCTCCGGGC 929
 OY 869 GAACCTGCCAGCTTGGTCTATGTGACACACACTCGCATATGCGACCCAGAAATCT 928
 DB 930 AACCTGTACACATGAGTGTGACAGACACCCCTCACACATATACAGAGACACGAAAT 989
 OY 929 ACCCAATGCTGCTTATATGAACTACTGACAGAAATCCAGATGCTGTGCGCTCTT 988
 DB 990 TCCCTCGAAAATTTGATGAAACTACTGCGCCGCAATCTGACGAAAMAAAGGCCCAT 1049
 OY 989 ATTGTATACGAGGATCCCGGTGTGAGGTGAGTACTGCAACCTGACGCAATGCTCAG 1048
 DB 1050 GTGTGCATACACACACGCAAGTGGGTGGAGTACTGTAAATACCTGCTT 1100
 OY 1049 ACGGAGAAAGGATCGCGTCCGCTCCGACTGTACCCCGGTTCGAAAGCTTGAAGGCTC 1108
 DB 1101 -----CCTGTGACTCTCTCCAGATATCCAG-----GAACAAT 1133
 OY 1109 CTTCCGAACAAGACGCACTGAGCAAAAGCGCTGGGGTGGAGAGTCTACATGCTATAG 1168
 DB 1134 TGGCTCCACAGCACACCTGAGCTAACCCCTGTGTCCAGAGCTGCTACATGTGTATG 1193
 OY 1169 GACAGAGTTATCGAGGACACTCTCCACACTGTCTGACAGAAAGAACTGCCAAGCTTGGT 1228
 DB 1194 GACAGAGTACCGAGGACACTCTCTCCACACACACAGAAAGAAAGTGTGCTTGGT 1253
 OY 1229 CATGTATGACACCACTGCAATAGTGGACCCCAAGTAATACCAATAGTGTGCTTGA 1288
 DB 1254 CATGTATGACACCACTGCAATAGTGGACCCCAAGTAATACCAATAGTGTGCTTGA 1313
 OY 1289 TCATGAACCTACTGAGGAATCCAGATGCTGTGACCTCTTATTGTATACGAGGATC 1348
 DB 1314 CAATGAACCTACTGAGGAATCCAGATGCGATTAAGGCCCTGTGTATTACACAGACC 1373
 OY 1349 CCGGTGTAGTGGAGTACTGCAATGCAATGCTCAAGCCAGAGAGGAGTGTGCG 1408
 DB 1374 CCAAGGTGTAGTGGAGTACTGCAATGCAATGCTCAAGCCAGAGAGGAGTGTG 1433
 OY 1409 TCGGCTCCGACTGTATACCCCGTTCCAGGCTTAGAGGCTCTTCCGAACAAGACCGA 1468
 DB 1434 TAGCACTCCGCTGTGTCTGCTTCCAGATGTAGAGACTCTTCCAGAGAA----- 1466
 OY 1469 CTGAGCAAAAGGCTGTGGGTCCAGAGTCTACCATGTTATGACAGATTATTCGAGGCA 1528

DB 1487 -----GACTGTATGTTGGGAATGGGAAAGATACCGAGCA 1523
 OY 1529 CATACTCCACCACTGTGACAGAAAGACCTCCAGAGTGTGCTATATGACACCACT 1588
 DB 1524 AGAGGCGCACCACTGTATATCTGGAGCGCCATCGCTGCGCCAGAGACCCCATTA 1583
 OY 1589 CCGATAG--TGGACCCCAAGTAATACCAATAGTGTGCTGTATGATCAATGAACTACTGA 1645
 DB 1584 GACACGCAATTTTACCTCCAGAGCAAAATCCAGGGGGGTCTGAAAAAATTAAGTCC 1643
 OY 1646 GGAATCCAGA--TGTGTGGACACTCTTATTGTATAGAGAGATCCCGGTGCAGGT 1702
 DB 1644 GTAACCTCGATGTGATGTAGGTGTCTCTGTGTCTACACGCAAAATCCAGAAATCTTT 1703
 OY 1703 GGAGTACTGCAACCTGACGCAATGCTCAGACGC 1736
 DB 1704 ACGACTAGTGTATGTCCCTCAGTGTGGGCCCC 1737

RESULT 9

AAQ40319
 ID AAQ40319 standard; cDNA: 2679 BP.

XX AAQ40319;

XX 17-AUG-1993 (first entry)

XX Sequence of a plasminogen cDNA.

XX Zymogen; fibrinolytic activity; cleavage; ss.

XX Synthetic.

XX Key Location/Qualifiers

FT CDS 22..2454

FT /tag= a

PN US5200340-A.

PD 06-APR-1993.

PF 22-MAY-1987; 87US-0053412.

PR 22-MAY-1987; 87US-0053412.

PA (ZYMO) ZYMOGENETICS INC.

PI Foster DC, Mulvihill ER, Ohara PJ, Pingel K, Yoshitake S;

DR WPI: 1993-133739/16.

DR P-PSDB: AAR34428.

XX Human tissue plasminogen activator single chain form fibrinolytic

PT agent - comprises thrombin cleavable zymogen stimulating amido

PT lytic activity, for lysing clots in heart attack and stroke

PT victims and suppressing fibrin matrix

XX Example: Fig 10A, 10B, 10C; 22pp: English.

XX A lambda phage clone comprising a cDNA sequence encoding

CC plasminogen was obtained from Dr. Mark Marten at the University of

CC Washington. The cDNA was isolated from a human liver library by

CC probing with the partial sequence of Malinowski et al. The sequence

CC of the complete cDNA and the encoded amino acid sequence are shown

CC in AAQ40319 and AAR34428.

XX Sequence 2679 BP; 750 A; 659 C; 675 G; 595 T; 0 other;

Query Match 4.7%; Score 236.6; DB 14; Length 2679;
 Best Local Similarity 58.6%; Pred. No. 3.7e-63;
 Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

OY 809 AGTGTACCATGTAATGACAGAGTTATGACAGACATCTCCACCACTGTCCAGAGAA 868

```

Db      842 ACTGCTGAGAGGAAACAGTGAACATATGCGGGAGATGTCGTTTACCGTCTCCGGGC 901
Qy      869 GAACCTGCCAAGCTTGGTTCATCTATGACACCACCTGGCATGTGCGACCCAGAAATACT 928
Db      902 AACACTGTCAGAGGTGAGGTGCACAGACCCCTCACACATACAGACACCCGAAAACT 961
Qy      929 ACCCAATGCTGCTTGCATGATGAACTACTGACAGATTCAGATGCTGTGGCAGCTCTT 988
Db      962 TCCCTCGCAAAATTTGGATGAACATACCTGCCCATCTGTGACGGAAAAAGGCCCAT 1021
Qy      989 ATTGTTATACGAGGATCCCGGTGTCAAGTGGAGTACTGCAACTGACGCAATGCTCAG 1048
Db      1022 GGTGCCATACAAACCAAGGCCAAGTGGGTGGAGTACTGTAAGATACCT----- 1072
Qy      1049 ACGCAGAAGGAGCTGCGTCCGCGCTCCGACTGTACCTCCGCGTTCACAGCTAGAGGCTC 1108
Db      1073 -----CCTGTGACTCTCTCCCGATCCACG-----GAACAAAT 1105
Qy      1109 CTTCCGAACACACCGACGTAGCAAAAGCCTGGGGTGCAGAGGTGCTACCATGTATATG 1168
Db      1106 TGGCTCCACACGACACCTAGCTAACCCCTGTGTCTCAGAGCTGCTACCATGTGATG 1165
Qy      1169 GACAGAGTTATGAGGACATACCTCCACACTGTACAGAGAAAGAACCTGCAAGCTGTGT 1228
Db      1166 GACAGAGCTACCGAGGACATCTCCACACACACAGAGAAAGAGTGTCACTTGTGT 1225
Qy      1229 CATCTATGACACACACTGCGATAGTCGAGCCCGACGAAATACCTACCAAAATGCTGGCTGA 1288
Db      1226 CATCTATGACACACACCGGACACGAAAGCCCGAAGAAATACCCAAATGCTGGCTGA 1285
Qy      1289 TCATGACACTGTGAGGAATCCAGATGCTGTGGCAGGCTCTTATTTATACGAGGATC 1348
Db      1286 CATGATACACTGTGAGGAATCCAGATGCTGTGGCAGGCTCTTATTTATACGAGGATC 1345
Qy      1349 CCGGTGTGAGTGGAGTACTGTCAACTGTGACGCAATGCTGACAGCAAGAGGACTGCCG 1408
Db      1346 CCAGCTCAGGTGGAGTACTGTCAACTGTGACGCAATGCTGACAGCAAGAGGACTGCCG 1405
Qy      1409 TCGCGCTCCGAGCTGTACCCCGGTTCACAGCTTAGAGGCTCTTCCGAAACAAGACCGA 1468
Db      1406 TAGCACTCCGCTGTGTGTGCTTCAGATGTAGAGACTCTCTCGAAGAA----- 1458
Qy      1469 CTGAGCAAAAGCCTGGGTGACAGAGTCTTACATGTTATGACAGAGTTATGAGGCA 1528
Db      1459 -----GACTGTATGTTTGGGAATGGGAAAGGATACCGAGGCA 1495
Qy      1529 CATACTCCACACTGTACACAGAGAAAGCTGCAAGGCTGGTCATGATGACACACACT 1588
Db      1496 AAGAGGCGACACTGTACTGTGAGCGCATGCTCCAGAGCTGGGCTGCCAGAGGCCCAT 1555
Qy      1589 CGCATAG---TCGAGCCCGAGATACTACCAAAATGCTGCTGTGATGATGAATACTGCA 1645
Db      1556 GACACAGCAATTTTCACTCCAGAGAACAAATCCAGCGCGGTCTGAAAAAAATTAAGTCC 1615
Qy      1646 GGAATCCAA---TGCTGGCAGCTCTTATTTATGATGAGGAGTCCCGGTGTGAGGT 1702
Db      1616 GTTAACTTATGTGTATGTAGTGTGCTGCTGTACACGCAAAATCCAAAGAACTTT 1675
Qy      1703 GGGAGTACTGCAACTGACGCAATGCTCAGAGC 1736
Db      1676 ACGACTACTGTATGTCTCTCAGATGTGGCGGCC 1709

```

RESULT 10

AAK35376 ID AAK35376 standard; DNA: 2732 BP.

XX AAK35376;

XX 16-JUL-1999 (first entry)

DE SEQ ID 51 of W0916889.

```

XX      XX      Angiostatin; endostatin; interferon; thrombospondin;
KW      KW      Interferon-inducible protein; platelet factor 4; anti-angiogenic;
KW      KW      anti-tumor; multifunctional protein; angiogenic-mediated disease;
KW      KW      cancer; diabetic retinopathy; macular degeneration; arthritis;
KW      KW      tumor cell production; ss.
OS      OS      Homo sapiens.
XX      XX      MO9916889-A1.
PN      PN      08-APR-1999.
XX      XX      30-SEP-1998; 98WO-US20464.
PF      PF      01-OCT-1997; 97US-0060609.
PR      PR      (SEAR ) SEARLE & CO G D.
PA      PA      Bolanowski MA, Caparon MH, Casperson GF, Gregory SA;
XX      XX      Klein BK, McKearn JP;
PI      PI      WPI; 1999-255098/21.
XX      XX      New multifunctional proteins useful for treating angiogenic-mediated
PT      PT      diseases
PS      PS      Disclosure: Page 85-86; 121pp; English.
XX      XX      The specification describes multifunctional proteins which comprise
XX      XX      combinations of angiostatin, endostatin, interferon, thrombospondin,
XX      XX      interferon-inducible protein and platelet factor 4, and have
XX      XX      anti-angiogenic and/or anti-tumor activity. The multifunctional protein
XX      XX      may exhibit useful properties such as having similar or greater
XX      XX      biological activity when compared to a single factor or by having
XX      XX      improved half-life or decreased adverse side effects, or a combination
XX      XX      of these properties. The proteins can be used for treating an
XX      XX      angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
XX      XX      degeneration, or arthritis. They can also be used for inhibiting the
XX      XX      production of tumor cells (characteristic of lung, breast, ovarian,
XX      XX      prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
XX      XX      hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
XX      XX      growth. The present sequence is used in the course of the invention.
SQ      SQ      Sequence 2732 BP; 757 A; 667 C; 690 G; 618 T; 0 other;
XX      XX
XX      XX      Query Match 4.7%; Score 236.6; DB 20; Length 2732;
XX      XX      Best Local Similarity 58.6%; Pred. No. 3.7e-63;
XX      XX      Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

```

```
QY 1169 GACAGAGTATGAGGACACATCTCAACCTGTGCACAGGAGAAGACCTGCAGCTTGGT 1228
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1199 GACACAGAGCTACCGAGGACGACATCTCCACACACACACAGAAAGAAAGTGTCACTTGGT 1258
QY 1229 CATCTATGACACACACATCTGCATAGTCGACCCCGAATACTACCAAAATGCTGGCTTTA 1288
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1259 CATCTATGACACACACACCGGACAGAAAGCCCGAATACTACCAAAATGCTGGCTTTA 1318
QY 1289 TCATGACTACTGACGAATCCAGATCTGTGGAGCTCTTATGTTATAGAGGATC 1348
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1319 CATGTAAGTACTGACGAGAAATCCAGATGCCATTAAGGCCCTCTGTATTACACAGACCC 1378
QY 1349 CCGGTGTAGGTGGAGTACTGCACACTGACGCAATGCTCAGACGAGAAAGGACCTGCCG 1408
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1379 TCAGGCTCAGGTGGAGTACTGCACAACTGAAATAAATCTCAGAAACGAGAGTGTGG 1438
QY 1409 TCGGCGCTCCGACTGTTCACCCGGTTCAGAGCCATAGAGGCTCTTCCGAAACAGACCGA 1468
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1439 TAGCACCTCCGCTGTCTGTCTCTCAGATGTAGAGACTCTTCCGAGAA----- 1491
QY 1469 CTGACCAAAAGGCTGGGGTGCAGAGTCTACCATGGTAATGGACAGAGTTATCGAGGCA 1528
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1492 -----GACTGTATGTTGGGAATGGAAAGATACCGAGGCA 1528
QY 1529 CATCTCCACCACTGTGCACAGGAAGACCTGCAGAGCTTGTATGACACCACT 1588
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1529 AAGAGGCGACCACTGTACTGTGACGCGCATGCCAGGCTGGCTCCGAGAGCCCATTA 1588
QY 1589 CGCATAG--TCGACCCCAAGAAATACCAAAATGCTGGCTGTATCATGAATCTGTGA 1645
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1589 GACACAGCATTTTCACTCTCAGAGCAAAATTCACGGGGGCTGTGAAAAAATTTACTGCC 1648
QY 1646 GGAATCCAGA--TGCTGTGACGCTCTTATGTTATAGAGGATCCCGGTGTCAAGT 1702
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1649 GTAACTCGATGTGATGTAGTGTGTGCTGTCTGTACAGCAAAATTCAGAAATTTT 1708
QY 1709 GGGAGTACTGCACCTGACGCAATGCTCAGAGCG 1736
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1709 ACGACTACTGTGATGTCCCTCAGTGTGCGGCC 1742

RESULT 11
ABN81696
ID ABN81696 standard; DNA: 2732 BP.
XX
AC ABN81696;
XX
DF 03-SEP-2002 (first entry)
XX
DE Human plasminogen encoding DNA.
XX
KW Human; plasminogen; lys plasminogen; glu plasminogen; A61; P22;
KW vascular endothelial; cell proliferation; anti-angiogenesis; cancer;
KW rheumatoid arthritis; Crohn's disease; diabetic retinopathy; anti-cancer;
KW cytostatic; gynaecological; neuroprotective; antiproliferative;
KW antiarteriosclerotic; dermatological; antineoplastic; antiarthritic;
KW antiinflammatory; antidiabetic; ophthalmological; immunosuppressive;
KW cardiant; vulnery; vasotropic; anti-tumour; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 1..2487
FT /*tag= a
FT /product= "plasminogen"
FT sig_peptide 55..111
FT /*tag= b
FT mat_peptide 112..2484
FT /*tag= c
FT /*product= "glu plasminogen"
FT mat_peptide 343..2484
FT /*tag= d
```

```
FT FT mat_peptide /product= "lys plasminogen"
FT 343..1515 /*tag= e
FT FT mat_peptide /product= "A61 isoform"
FT 343..1524 /*tag= f
FT FT mat_peptide /product= "A61 isoform"
FT 343..651 /*tag= g
FT FT mat_peptide /product= "P22"
XX
XX WO200244328-A2.
XX
XX 06-JUN-2002.
XX
XX 28-NOV-2001; 2001WO-US44515.
XX
XX 28-NOV-2000; 2000US-253725P.
XX
XX (WAIS/) WAISMAN D M.
XX (KASS/) KASSAM G.
XX (KWON/) KWON M.
XX
XX Waisman DM, Kassam G, Kwon M;
XX
XX WPI: 2002-52706/56.
XX P-PSDB: ABB83795.
XX
XX Novel naturally occurring fragment A16 or P22 of plasminogen, useful
XX for treating for anti-angiogenic treatment of a mammal suffering from
XX cancer and inhibiting proliferation of vascular endothelial cells
XX
XX Claim 26; Fig 2; 88pp; English.
XX
XX The invention relates to an isolated naturally occurring fragment A61
XX or p22 polypeptide (1) of plasminogen or an isolated polypeptide
XX comprising 103, 391 or 394 contiguous amino acids of amino terminal of
XX plasminogen (ABB83795). (1) has vascular endothelial cell proliferation
XX inhibitor activity. (1) is useful for anti-angiogenic treatment of a
XX mammal suffering from cancer, acoustic neuromas, neurofibromas,
XX trachomas, pyogenic granulomas, telangiectasias, psoriasis, scleroderma,
XX atherosclerosis, rheumatoid arthritis, Crohn's disease, endometriosis,
XX adiposity, pyogenic granuloma, rubeosis, and diabetic retinopathy,
XX retinopathy of prematurity, neovascular glaucoma, retrolental
XX fibroplasia, graft rejection, myocardial angiogenesis, plaque
XX neovascularisation, haemophilic joints, angiodiroma and wound
XX granulation and inhibiting proliferation of vascular endothelial cells,
XX in combination with an anti-cancer agent e.g. methotrexate, mitozantrone,
XX paclitaxel, vinblastine, 5-fluorouracil, cisplatin, leucovorin,
XX cyclophosphamide and oncolytic virus.
XX
XX Sequence 2732 BP; 758 A; 669 C; 688 G; 617 T; 0 other:
XX
XX Query Match 4.7%; Score 236.6; DB 24; Length 2732;
XX Best Local Similarity 58.6%; Pred. No. 3,7e-63;
XX Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;
QY 809 AGGTGTACAGATGTATGAGACAGGATTTGAGGACATCTCCACACCTGTGCAGCGAA 868
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 875 AGTGTGTGAAGGAACAGGATGAAACTATCGGGGAATGTGGCTGTACCGTTTCCGGGC 934
QY 869 GAACCTGCCAAGCTTGTGATCTATGACACCACTGCATGTGAGACCCAGAACTACT 928
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 935 ACACCTGTGAGCAGCTGAGTGCACAGACCCCTCACAACATACAGACACCGAAAACT 994
QY 929 ACCCAAAATGCTGTGATCATGAACTACGACAGATCCAGATGTGCGGACGCTCTT 988
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 995 TCCTCGCAAAATTTGGATGAAATCTATGCTCCGATCTGACGAAAAAGGCCCAT 1054
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 989 ATTGTTATCGAGGAGATCCCGGTGTGAGTGTGAGTACTCAACCTGACGCAATGCTAG 1048
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1055 GGTGCATACCAACCAACAGCAAGTGGGTGGAGTGTAGATACCGT----- 1105
```


Db 1728 GTGATGTCCTCAGTGTGGGCCCC 1752

RESULT 13

ID AAQ12542 standard; DNA; 2756 BP.

AC AAQ12542;

XX 23-SEP-1991 (first entry)

XX Encodes Plasminogen muteln X1 with factor Xa cleavage site.

XX protease; fibrinolysis; blood clotting; ss.

XX Key Location/Qualifiers

FT mutation /tag= a

FT CDS /note= "replaces CCT (Pro codon) with two codons"

FT /tag= b

FT /product= modified plasminogen

PN MO9109118-A.

XX 27-JUN-1991.

XX 07-DEC-1990; 90WO-G001912.

XX 07-DEC-1989; 89GB-0027722.

XX 07-DEC-1990; 90WO-GB01911.

XX (BRBI-) BRIT BIO-TECHN LTD.

PI Dawson KM, Edwards RM, Forman JM;

DR WPI: 1991-208145/28.

XX P-PSDB: AAR12933.

PT Activatable fibrinolytic and antithrombic proteins - activated by

XX e.g. factor Xa, thrombin or activated protein C

XX Claim 16; Fig 2 and Fig 4; 73pp; English.

XX This sequence encodes a plasminogen mutant cleavable by Factor Xa.

XX Activation is localised to the thrombus because cleavage to plasmin

XX is by an enzyme of the blood clotting pathway. Compositions

XX comprising the mutant plasminogen encoded by this sequence are used

XX for treatment or prevention of thrombosis, etc.

XX See AAQ12543-Q12558.

XX Sequence 2756 BP; 767 A; 670 C; 697 G; 622 T; 0 other;

XX Query Match 4.7%; Score 236.6; DB 12; Length 2756;

XX Best Local Similarity 65.6%; Pred. No. 3.7e-63;

XX Matches 410; Conservative 0; Mismatches 179; Indels 36; Gaps 3;

QY 92 CAGAGAGACCTGAGCAAGCATGTGTCACAGAGACCTGCCAAGCTTGTCATCTATGA 211

Db 1158 CAGACACCACTGAGCTAACCCCTGTGTCCAGAGACTGCCATGTCATGACAGACT 1217

QY 152 ATCAGGACGACTCTCCACACTGTGCACAGAGACCTGCCAAGCTTGTCATCTATGA 211

Db 1218 ACCGAGGACATCTCTCCACCAACAGAGAAAGTGTCAAGCTTGTCATCTATGA 1277

QY 212 CACCAATCAACATTAATAGACACAGAAACTACCAATCTGCTTGATCATGAAC 271

Db 1278 CACCAACACCGGACACAGAAACCCACAAATCTGCTTGATCATGAAC 1337

QY 272 ACTCAGAGAAATCCAGATGCTGTGACAGCTCTTATTGTATATGAGAGATCCGGTGCA 331

Db 1338 ACTCAGAGAAATCCAGATGCTGTGACAGCTCTTATTGTATATGAGAGATCCGGTGCA 1397

QY 332 GGTGGAGTACTGCAACCTGACGCAATGCTCAGACGAGAGGAGACTGCCGTGGCCCTC 391

Db 1398 GGTGGAGTACTGCAACCTGAGAAAATGCTCAGAGAAACAGAGAGTGTGATGACACTC 1457

QY 392 CGACTGTTACCCCGTTCCAGACCTAGAGGCTCTTCGAGACAGCCGACTGAGCAAA 451

Db 1458 CGCTGTTGTCCTGCTCCAGATGTAGAGACTCTTCGAGAA----- 1501

QY 452 GGCTGGGGTGCAGAGTGCATCATATGAGACAGATGTCAGGACATACACTCA 511

Db 1502 -----GACTGATGTTTGGAAATGGAAAGGATACGAGCAAGAGCGCA 1547

QY 512 CCAGTGTACAGAGAAAGCAACCTGTCATCTATGACACACACTGCATAG-- 569

Db 1548 CCAGTGTACTGGAGCGCAGCAGAGACTGGCTGCCAGGCCCCATTAGACACACA 1607

QY 570 -TCGAGACCCAGAAATACCAACAAATGCTGTGATGATGAACTTCGAGATTCAG 628

Db 1608 TTTTCACTCCAGAGCAAAATCCAGGGGCTGTGGAAAAATTATGCTGTAACCTG 1667

QY 629 A---TCTGTGGCAGCTCTTATTGTATATGAGAGGATCCCGGTGTCAGGTGGAAGTACT 685

Db 1668 ATGATGATGTAGTGTGCTGCTGCTGCTACGACAAATCCAAAGAAATCTTACACTACT 1727

QY 686 GCAACCTGACGCAATGCTCAGAACCC 710

Db 1728 GTGATGTCCTCAGTGTGGGCCCC 1752

RESULT 14

ID AAQ12548 standard; DNA; 2756 BP.

XX AAQ12548;

XX 23-SEP-1991 (first entry)

XX Encodes Plasminogen muteln T2 with thrombin cleavage site.

XX protease; fibrinolysis; blood clotting; ss.

XX Key Location/Qualifiers

FT mutation /tag= a

FT /note= "replaces CCTGCA (Pro-Gly) with GGTGTCCA

FT CDS 65..2500

FT /tag= b

FT /product= modified plasminogen

PN MO9109118-A.

XX 27-JUN-1991.

XX 07-DEC-1990; 90WO-G001912.

XX 07-DEC-1989; 89GB-0027722.

XX 07-DEC-1990; 90WO-GB01911.

XX (BRBI-) BRIT BIO-TECHN LTD.

PI Dawson KM, Edwards RM, Forman JM;

DR WPI: 1991-208145/28.

XX P-PSDB: AAR12939.

PT Activatable fibrinolytic and antithrombic proteins - activated by

XX e.g. factor Xa, thrombin or activated protein C

XX Claim 16; Fig 2 and Fig 5; 73pp; English.

XX This sequence encodes a plasminogen mutant cleavable by thrombin.

XX Activation is localised to the thrombus because cleavage to plasmin

CC is by an enzyme of the blood clotting pathway. Compositions comprising the mutant plasminogen encoded by this sequence are used CC for treatment or prevention of thrombosis, etc.
CC See AA012542-Q12558.

XX Sequence 2756 BP; 765 A; 671 C; 697 G; 623 T; 0 other;

Query Match 4.7%; Score 236.6; DB 12; Length 2756;
Best Local Similarity 65.6%; Pred. No. 3.7e-63;
Matches 410; Conservative 0; Mismatches 179; Indels 36; Gaps 3;

```
OY 92 CAGCAGCAGCTGACGAAAGCAGTGGTCCAGATGCTTACCACTGGTGTGACAGAGTT 151
DB 1138 CAGCAGCAGCTGACGAAAGCAGTGGTCCAGATGCTTACCACTGGTGTGACAGAGTT 1217
OY 152 ATCGAGGACAGCTACTCCAGCACTGTCCAGAGAGAGACTGCCAAGCTTGCTATGTA 211
DB 1218 ACCGAGGACATCTCCAGCCACACAGAGAAAGTGTCACTGCTGATCATGTA 1277
OY 212 CACCATCAACATATATAGGACACAAATACCAATGCTGGCTGATCATGAGT 271
DB 1278 CACCATCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1337
OY 272 ACTGAGGAAATCAGATGCTGTGACAGCTCTTATTTATACGAGGAGTCCGGGTGCA 331
DB 1338 ACTGAGGAAATCAGATGCTGTGACAGCTCTTATTTATACGAGGAGTCCGGGTGCA 1397
OY 332 GGTGGGATGACTGACAGCTGACAGCAATGCTCAACGACGAGAGTCCGCGGCTC 391
DB 1398 GGTGGGATGACTGACAGCTGACAGCAATGCTCAACGAGAGAGTCTGTGACAGCTC 1457
OY 392 CGACTGTACCCGCTTCCAGAGCTGACAGCTCTTCCGAAACAGAGCAGCTGAGCAAA 451
DB 1458 CGCTGTGCTGCTTCCAGATGTAGAGCTCTTCCGAAAGAGAGAGAGAGAGAGAGAG 1501
OY 452 GGCCTGGGCTGACAGAGTGTACCATGTATGACAGAGTATTCAGAGCAGATATCCA 511
DB 1502 -----GACTGTATGTTTGGAGATGGAGAAAGATACCCAGAGAGAGAGAGAGAG 1547
OY 512 CCAGTGTACAGGAAAGCCTGCAAGCTTGATCATATGACACGACAGCTGCATAG-- 569
DB 1548 CCAGTGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1607
OY 570 -TCGAGCCCAAGATATCTACCAATGCTGTGATCATGATGATCTGAGAGATCCAG 628
DB 1608 TTTTCACTCCAGAGCAAAATCCAGGAGGCTGTGAAAAAAATTAATCTGCTTAACCTG 1667
OY 629 A---TGCTGTGAGAGCTCTTATTTATGAGAGAGATCCCGGTGTCAGGTGGAGTACT 685
DB 1668 ATGCTGTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1727
OY 686 GCAACCTGAGCAATGCTCAGAGCG 710
DB 1728 GTGATGTCTCTCAGTGTGCGGCC 1752
```

RESULT 15

AA012552
ID AA012552 standard; DNA; 2756 BP.

XX AA012552;

XX 23-SEP-1991 (first entry)

XX Encodes plasminogen mutuin T13 with thrombin cleavage site.

XX protease; fibrinolysis; blood clotting; ss.

XX Key Location/Qualifiers

FT mutation

1796..1804

FT /tag= a

FT /note= "replaces CCTGGA (i.e. two codons)"

```
PN WO9109118-A.
XX 27-JUN-1991.
PD 07-DEC-1990; 90WO-G001912.
XX 07-DEC-1989; 89GB-0027722.
PR 07-DEC-1990; 90WO-GB01911.
XX (BRB-) BRIT BIO-TECHN LTD.
XX Dawson KM, Edwards RM, Forman JM;
XX WPI, 1991-208145/28.
XX P-PDB; AARI2943.
XX
XX Claim 16; Fig 2 and Fig 5; 73pp; English.
XX
XX This sequence encodes a plasminogen mutant cleavable by thrombin.
XX Activation is localised to the thrombus because cleavage to plasmin
XX is by an enzyme of the blood clotting pathway. Compositions
XX comprising the mutant plasminogen encoded by this sequence are used
XX for treatment or prevention of thrombosis, etc.
XX See AA012542-Q12558.
XX
XX Sequence 2756 BP; 765 A; 671 C; 695 G; 625 T; 0 other;
SQ
Query Match 4.7%; Score 236.6; DB 12; Length 2756;
Best Local Similarity 65.6%; Pred. No. 3.7e-63;
Matches 410; Conservative 0; Mismatches 179; Indels 36; Gaps 3;
OY 92 CAGCAGCAGCTGACGAAAGCAGTGGTCCAGATGCTTACCACTGGTGTGACAGAGTT 151
DB 1138 CAGCAGCAGCTGACGAAAGCAGTGGTCCAGATGCTTACCACTGGTGTGACAGAGTT 1217
OY 152 ATCGAGGACAGCTACTCCAGCACTGTCCAGAGAGAGACTGCCAAGCTTGCTATGTA 211
DB 1218 ACCGAGGACATCTCCAGCCACACAGAGAAAGTGTCACTGCTGATCATGTA 1277
OY 212 CACCATCAACATATATATGAGACACAGAAATCTACCAATGCTGTGATCATGTA 271
DB 1278 CACCATCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1337
OY 272 ACTGAGGAAATCAGATGCTGTGACAGCTCTTATTTATACGAGGAGTCCGGGTGCA 331
DB 1338 ACTGAGGAAATCAGATGCTGTGACAGCTCTTATTTATACGAGGAGTCCGGGTGCA 1397
OY 332 GGTGGGATGACTGACAGCTGACAGCAATGCTCAACGAGAGAGTCCGCGGCTC 391
DB 1398 GGTGGGATGACTGACAGCTGACAGCAATGCTCAACGAGAGAGTCTGTGACAGCTC 1457
OY 392 CGACTGTACCCGCTTCCAGAGCTGACAGCTCTTCCGAAACAGAGCAGCTGAGCAAA 451
DB 1458 CGCTGTGCTGCTTCCAGATGTAGAGACTCTTCCGAAAGAGAGAGAGAGAGAGAGAG 1501
OY 452 GGCCTGGGCTGACAGAGTGTACCATGTATGACAGAGTATTCAGGCAATATCCA 511
DB 1502 -----GACTGTATGTTTGGAGATGGAGAAAGATACCGAGAGAGAGAGAGAGAG 1547
OY 512 CCAGTGTACAGGAAAGCCTGCAAGCTTGATCATATGACACGACAGCTGCATAG-- 569
DB 1548 CCAGTGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1607
OY 570 -TCGAGCCCAAGATATCTACCAATGCTGTGATCATGATGATCTGAGAGATCCAG 628
DB 1608 TTTTCACTCCAGAGCAAAATCCAGGAGGCTGTGAAAAAAATTAATCTGCTTAACCTG 1667
OY 629 A---TGCTGTGAGAGCTCTTATTTATGAGAGAGATCCCGGTGTCAGGTGGAGTACT 685
DB 1668 ATGCTGTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1727
```

Oy 686 GCAACCTGACGCAATGCTCAGACGC 710
| | | | | | | | | |
Db 1728 GTGATGTCCCTCAGTGTGCGGCCCC 1752

Search completed: March 6, 2003, 05:12:19
Job time : 932 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 06:19:18 ; Search time 148 Seconds
(without alignments)
10360.701 Million cell updates/sec

Title: US-09-923-515-3_COPY_1_5000

Perfect score: 5000
Sequence: 1 ctggagattggagacactt.....caagcttgatcatatgac 5000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	236.6	4.7	2497	1	US-08-643-219-12
2	236.6	4.7	2497	2	US-09-131-995-12
3	236.6	4.7	2497	2	US-08-832-087B-12
4	236.6	4.7	2497	3	US-08-851-350-12
5	236.6	4.7	2497	4	US-09-132-154-12
6	236.6	4.7	2679	6	5200340-7
7	236.6	4.7	2753	1	US-07-854-603-1
8	235	4.7	2296	1	US-07-750-080A-18
9	235	4.7	2296	3	US-08-651-472-18
10	235	4.7	2296	4	US-08-358-928-18
11	216.2	4.3	1134	4	US-09-206-059-29
12	92.4	1.8	2542	1	US-08-441-370-1
13	67.6	1.4	645	3	US-08-985-526-22
14	67.6	1.4	1284	3	US-08-985-526-24
15	62.8	1.3	2172	4	US-08-030-410-2
16	62.8	1.3	2184	4	US-07-815-333A-1
17	62.8	1.3	2288	1	US-08-290-937B-4
18	62.8	1.3	2288	1	US-07-838-410-2
19	59.2	1.2	2188	1	US-07-882-925A-4
20	59.2	1.2	2188	1	US-08-184-012C-4
21	55.2	1.1	315	1	US-08-242-035A-11
22	54.8	1.1	2021	1	US-07-882-925A-3
23	54.8	1.1	2021	1	US-08-184-012C-3
24	54.8	1.1	2216	2	US-08-666-082B-2
25	54.8	1.1	2219	1	US-07-882-925A-1
26	54.8	1.1	2219	1	US-07-882-925A-2
27	54.8	1.1	2219	1	US-08-184-012C-1

28	54.8	1.1	2219	1	US-08-184-012C-2	Sequence 2, Appl
29	54.8	1.1	2232	1	US-08-334-177-1	Sequence 1, Appl
30	54.8	1.1	2232	5	PCT-US95-13830-1	Sequence 1, Appl
31	54.8	1.1	2262	1	US-07-882-925A-7	Sequence 7, Appl
32	54.8	1.1	2262	1	US-08-184-012C-7	Sequence 7, Appl
33	51.4	1.0	3398	5	PCT-US95-08493-12	Sequence 12, Appl
34	35.6	0.7	6100	1	US-07-882-925A-6	Sequence 6, Appl
35	35.6	0.7	6100	1	US-08-184-012C-6	Sequence 6, Appl
36	33.4	0.7	6751	1	US-07-882-925A-5	Sequence 5, Appl
37	33.4	0.7	6751	1	US-08-184-012C-5	Sequence 5, Appl
38	32.6	0.7	4092	2	US-08-469-537A-106	Sequence 106, App
39	32.2	0.6	705	4	US-08-998-416-1136	Sequence 1136, Ap
40	32	0.6	32	1	US-07-832-905B-19	Sequence 19, Appl
41	32	0.6	32	2	US-08-700-757-19	Sequence 19, Appl
42	31.8	0.6	4626	1	US-08-306-691B-22	Sequence 22, Appl
43	31.8	0.6	4626	5	PCT-US93-06251-27	Sequence 27, Appl
44	31.6	0.6	1869	3	US-08-952-967-7	Sequence 7, Appl
45	30.8	0.6	2058	4	US-09-560-639-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-643-219-12
; Sequence 12, Application US/08643219
; Patent No. 5801146
; GENERAL INFORMATION:
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/643, 219
; FILING DATE: 06-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Dianne
; REGISTRATION NUMBER: 40, 943
; REFERENCE/DOCKET NUMBER: 5940, US, P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; TELEFAX: 847-938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2497 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-643-219-12

Query Match 4.7%; Score 236.6; DB 1; Length 2497;
Best local Similarity 58.6%; Pred. No. 1.3e-70;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;
QY 809 AGTCATCATGTATGACAGAGTTATCGAGGCACATCTCCACCACTCTCAGAGAA 866

```

Db      870 AGTGTCTGAGAGGAACAGGTGAAACATATGCGGGATGmGGCTTTAAGGTGTCCGGGC 929
      869 GAACCTGCCAAGCTTGGTGCATCTATGACACCACTGGCTAGTGGACCCCAATACT 928
      930 AACCTGTGAGAGTGGAGTGCACAGCCCTCAGACATTAACAGACACCCGAAACT 989
      929 ACCCAATGCTGGCTTGCATGAACTACTGACAGAAATCCAGATGCTGGACATCTT 988
      990 TCCCTGCAAAAATTTGATGAAAACATACGCCCATCTGACGAAAAAGGCCCAT 1049
      989 ATTGTATATACAGAGGATCCCGGTGCAGGTGGAGTACTCAACCTGACGCAATGCTCAG 1048
      1050 GGTGCAATACCAACCAACGCCAATGGGTGGAGTACTGTAAAGTACGCT----- 1100
      1049 ACGCAGAAAGGACTGCGCTCGGCTCCGACTGTACCCGGTTCAGAGCTAGAGGCTC 1108
      1101 -----CCTGTGACTCTCTCCCAAGTATCCAGC-----GAACAT 1133
      1109 CTTCGCAACAGCAGCAGCTAGCAAAAGCCCTGGGTGACAGAGTGTCTACATGTATG 1168
      1134 TGGCTCCACAGACACACCTAGGTAAACCTGTGGTCCAGAGCTCTACATGTATG 1193
      1169 GACAGATTATGAGGACATCTCTCCACACTGTCAAGAGAAAGACCTGCAAGCTTGGT 1228
      1194 GACAGAGCTACGAGGACATCTCTCCACACACAGAAAGAAAGTGTCTTGGT 1253
      1229 CATCTATGACACACATCTCGCATGTGGACCCAGAAATCTACCAATGCTGCTTGA 1288
      1254 CATCTATGACACACACCGGACCCAGAAAGCCCAAAATCTCCAAATGCTGCTTGA 1313
      1289 TCATGAACTAGTGCAGAAATCAGATCTGTGGAGCTCTTATGTTATGACAGGATC 1348
      1314 CAATGAACTAGTGCAGAAATCAGATCTGTGGAGCTCTTATGTTATGACAGGATC 1373
      1349 CCGGTGTAGTGGAGTACTGCAACTGAGCAATGCTCAAGAGCAAGAGGAGTCCG 1408
      1374 CAGAGGTAGTGGAGTACTGCAACTGAGCAATGCTCAAGAGCAAGAGGAGTCCG 1433
      1409 TCGGCTCTCAGTGTACCCGGGTTCAGAGCTAGAGGCTCTTCCGAAGACACGCA 1468
      1434 TAGCACTCTCGGCTGTGTCTCTGCTTCAGATGTAAGAGACTCTTCCGAAGAA----- 1486
      1469 CTGACAAAGAGCTGGGGTGCAGAGTGTCAATGTAAGAGAGTATGAGAGCA 1528
      1487 -----GACTGTATGTTGGGATGGGAAAGATACCGAGCA 1523
      1529 CATCTCCACACTGTCTACAGAAAGCTGCCAAGCTGGTATGATGACACCAACT 1588
      1524 AGAGGCGACACATCTTACTGTGAGAGCCATGCCAGACTGGCTGCCAGAGCCCAT 1583
      1589 GCGATAG--TGGACCCCAAGATATCAACCAATGCTGGCTTGTATGATGAACTATGCA 1645
      1584 GACACAGCATTTTCACTCCAGAGACAATATCACGGGGGGTGTGAAAAAATTTACTGCC 1643
      1646 GGAATCCAGA---TGCATGGCAGCTCTTATGTAATACAGAGATCCCGGTGTCAGGT 1702
      1644 GTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1703
      1703 GGGAGTACTGCAACCTGACGCAATGCTCAGAGC 1736
      1704 ACGACTACTGTGATGTCCCTCAGTGTGCGGCC 1737

```

RESULT 2
US-09-131-995-12

; Sequence 12, Application US/09131995

; Patent No. 5972896

; GENERAL INFORMATION:

; APPLICANT: Davidson, Donald J.

; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES

; AND METHODS FOR INHIBITING ANGIOGENESIS

; NUMBER OF SEQUENCES: 14

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/131,995
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/832,087
FILING DATE: 03-APR-1997
APPLICATION NUMBER: 08/643,219
FILING DATE: 06-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 5940.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
TELEFAX: 847-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2497 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-131-995-12

Query Match      4.7%   Score 236.6; DB 2:   Length 2497;
Best Local Similarity 58.6%   Pred. No. 1.3e-70;
Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

```


RESULT 4

US-08-851-350-12

; Sequence 12, Application US/08851350

; Patent No. 6057122

GENERAL INFORMATION:

APPLICANT: Abbott Laboratories

TITLE OF INVENTION: NOVEL ANTITUMORIGENIC PEPTIDES,

TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME AND METHODS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/851,350

FILING DATE: 05-MAY-1997

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Casuto, Dianne

REGISTRATION NUMBER: 40,943

REFERENCE/DOCKET NUMBER: 5940.US.P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847-938-3137

TELEFAX: 847-938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 2497 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-851-350-12

Query Match

Best Local Similarity 58.6%; Score 236.6; DB 3; Length 2497;

Matches 547; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

```
QY 809 AGTGTACCATGTATGACAGAGTTATGAGGACATACCTGACACCTGTCACAGGAA 868
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 870 AGTGTGTGAAGGAAGGAGGAAACATATCGCGGAATGTGCTGTCGGTCCGGGC 929
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 869 GAACCTGCCAAGCTTGTCTATGACACACACACTGCGATAGTGGACCCGAAATACT 928
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 930 ACACCTGTACAGCTGAGTGCACAGACCCCTCACACATTAACAGACAGCAAGAACT 989
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 929 ACCCAATGCTGCTGATCATGACATGACAGCAAGCAATGCTGTGGAGCTCTT 988
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 990 TCCCTCGCAAAATTTGATGAAACACTGCGCGCAATCTGTACGGAAGAAAGGCCCCAT 1049
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 989 ATTGTATACGAGGATCCGGGTGTCAGGTGGAGTACTGCAACCTGACGCAATGCTCAG 1048
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1050 GGTGCAATACCAACAGCAAGTGGGTGGAGTACTGTAAGATACCGT----- 1100
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1049 ACGGAGAGGAGACTGCGGTGCGCTCCGCTTACCCCGGTCCAGAGCCATGAGGCTC 1108
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1101 -----CCTGTGACTCTCTCCAGATATCCAG-----GACACAT 1133
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1109 CTTCCGAACAGCAGCATGAGCAAGGCTTGGGTGAGGAGTACCATGTGTAATG 1168
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
DB 1134 TGGCTCCACAGCACACCTGAGCTAACCCCTGTGTGTCAGGAGCTGTACCATGCTATG 1193
QY 1169 GACAGATTATCGAGGCACATCTACTCCACACATCTGTACAGGAAAGAACTGCCAGCTTGT 1228
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1194 GACAGGCTTACCGAGGACATCTCTCCACACACACACAGGAAAGAAAGTGAAGTCTTGGT 1253
QY 1229 CATCTATGACACACACTGCTCATAGTTCGACCCCAAGTAATACCAATATGCTGCTTGA 1288
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1254 CATCTATGACACACACACGCGGACCCAGAGAGCCCAAGAAACATACCAATATGCTGCTTGA 1313
QY 1289 TCATGACTACTGAGCAATTCAGATGCTGTGGAGCTCTTATTTATATACAGGAGATC 1348
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1314 CAATGAACTACTGACGAAATTCAGATGCGGATTAAGGCCCCGTGTTTATCACAGAGCC 1373
QY 1349 CCGGTGTAGGTGGAGTACTGCACTGACGCAATGCTCAGACGAGAGAGGAGCTGCG 1408
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1374 CCAGGTCAGTGGGAGTACTGCACTGAAATGCTCAGAGAACAGACGAGTGTG 1433
QY 1409 TCGGCTCGAGCTGTATTACCCCGGTTCCAGGCTAGAGGCTCTCTCCAGACAGACCGA 1468
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1434 TAGCAGCTCCGCTGTGTCTCTCCAGATGATGACAGCTCTCCGAGAGAA----- 1486
QY 1469 CTGACCAAGGCTGTGGGTGACAGAGTGCATGATGATGATGATGATGATGATGATGATG 1528
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1487 -----GACTGTATGTTTGGGAATGGGAATGGAAGATACGAGAGCA 1523
QY 1529 CATCTCCACACATGTCACAGGAAAGCTGCCAGCTTGTGATCTATGACACACACT 1588
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1524 AGAGGCGACACAGCTTACTGAGGACGCGCATGCGGAGCTGCGGCTCCAGAGAGCCATA 1583
QY 1589 CGCATAG---TCGAGCCCAAGTAATACCAATGCTGCTGTGATCATGATGATGATGATG 1645
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1584 GACAGAGATTTTCACTCCAGAGCAAAATCCAGGCGGCTGTGGAAGAAATTAATGCTCC 1643
QY 1646 GGAATCCAGA---TCTGTGCGAGCTCTTATTTATATACAGAGGATCCCGGTGCAAGT 1702
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1644 GTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1703
QY 1703 GGAAGTACTGCAACCTGACGCAATGCTCAGAGCG 1736
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1704 ACGACTACTGTGATGTCCCTCAGATGTCGCGGCC 1737
```

RESULT 5

US-09-132-154-12

; Sequence 12, Application US/09132154

; Patent No. 6251867

GENERAL INFORMATION:

APPLICANT: Davidson, Donald J.

TITLE OF INVENTION: NOVEL ANTITUMORIGENIC PEPTIDES

TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/132,154

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/832,087

FILING DATE: 03-APR-1997

APPLICATION NUMBER: 08/643,219

FILING DATE: 06-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 3940, US, P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
TELEFAX: 847-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2497 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-132-154-12

Query Match 4.7%; Score 236.6; DB 4; Length 2497;
Best Local Similarity 58.6%; Pred. No. 1.3e-70;
Matches 347; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

QY 809 AGTGTACCATGTTATGAGACAGAGTTATCGAGCACAATCTCCACCACCTGTGCAGAGAA 868
DB 870 AGTGTGTAAGGGAACAGAGTGAACCTATCGGGGAATGTGCTTAACGGTGTCCGGGC 929
QY 869 GAACCTGCCAAGCTTGTATCTATGACACCACTGCATATGTCGAGCCCGAGAACTCT 928
DB 930 ACACCTGTACAGCTGAGTGCACAGACCCCTCACACATTAACAGAGACACCGAATAACT 989
QY 929 ACCCAATCTGCTTGATCATGAACTACTGCAGAAATCCAGATGCTGTGGACCTCTT 988
DB 990 TCCCTGCAAAAATTGTGATGAATACTGTGCGCAATCTGACGAGAAAAAGGCCCAT 1049
QY 989 ATGTGTATACGAGGATCCCGGTGTACGTGGAGTACTGCAAGCTTACCACTGACCAATGCTCAG 1048
DB 1050 GGTGCATTAACCAACCAACCAAGCCAGTGGGTGAGTACTGTAAATACCGT----- 1100
QY 1049 ACAGAGAAAGGACTGCGCTCGGCTCCGACTGTTACCCCGTTCAGACCTAGAGCTC 1108
DB 1101 -----CCTGTGACTCTCTCCCACTATCCAGC---GAACAAT 1133
QY 1109 CTTCGGAACAGACGACCTGAGCAAGAGCCGGGGTGCAGAGAGTACCATGTTAATG 1168
DB 1134 TGGCTCCACAGCACCACCTGAGCTAACCCTGTGTGTCCAGAGACTGCTACCATGTGTATG 1193
QY 1169 GACAGAGTTATGAGGACACATCTCCACACTGTCAAGAAAGAACTGCCAAGCTTGT 1228
DB 1194 GACAGAGTTATGAGGACACATCTCCACACTGTCAAGAAAGAAAGTGTGAGTCTGTGT 1253
QY 1229 CATCTATGACACGACCTGCAATGTCGAGCCCGCAATTAATCCCAATGCTGTGTA 1288
DB 1254 CATCTATGACACGACCGGACCAAGAAAGCCCAAGAAATCCCAATGCTGTGTA 1313
QY 1289 TCATGAATCTGACAGAAATCCAGATGCTGTGGACCTCTTATTTATTCAGAGGATC 1348
DB 1314 CAATGAATCTGACAGAAATCCAGATGCTGTGGACCTCTTATTTATTCAGAGGATC 1373
QY 1349 CCGGTGTACAGTGGAGTACTGCAACCTGACAGCAATGCTCAGAGCAGAGAGGACTCCG 1408
DB 1374 CCAGGCTCAGGTGAGTCTGCAACTGAAATAATGCTCAAGAAAGAAAGCGGTGTG 1433
QY 1409 TCAGGCTCAGGCTGTTACCCCGGTTCCAAAGCCTAGAGGCTCTTTCGAAACAGACACGA 1468
DB 1434 TAGACACTCGGCTGTGTGTCTGTTCAGAGATGTAGAACTCTTCCGAGAA----- 1486
QY 1469 CTGAGCAAAAGCGCTGGGCTGCGAGAGTGTACATGTAATGAGACAGATTAATGAGGCA 1528
DB 1487 -----GACTGTATGTTTGGAAATGGAAGAGATCCGAGGCA 1523
QY 1529 CATACTCCACACTGTACAGAGAAAGAACTCCCAAGCTTGTCAATTATGACACAGACT 1588
DB 1524 AGAGGGGAGACACTGTACTGTGGAGCGCAATGTCAGAGACTGGGCTGCCAGAGCCCATTA 1583

QY 1589 CGCATAG---TCGACCCCAACCACTACTACCAATGCTGGTGTGATGACTACTGCA 1645
DB 1584 GACACACATTTTCACTCCAGAGACAAATCCACAGGCGGGTCTGAAAAAATTAATCTGCC 1643
QY 1646 GGAATCCAGA---TCGTGTGGCAGCTCTTATTTATTAACGAGGATCCCGTGTCAAGGT 1702
DB 1644 GTAACCTGTATGTGTATGTGTGGTGTCCCTGTGTACACGCAATTCAGAAAAACTTT 1703
QY 1703 GGGAGTACTGCAACCTGACGCAATGCTCAGAGCC 1736
DB 1704 ACAGACTACTGTGATGTCCCTCAAGTGTGGCGCCC 1737

RESULT 6

5200340-7
PATENT NO. 5200340
APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
ACTIVATORS
NUMBER OF SEQUENCES: 34
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/53,412
FILING DATE: 22-MAY-1987
SEQ ID NO: 7:
LENGTH: 2679
5200340-7

Query Match 4.7%; Score 236.6; DB 6; Length 2679;
Best Local Similarity 58.6%; Pred. No. 1.3e-70;
Matches 347; Conservative 0; Mismatches 315; Indels 72; Gaps 5;

QY 809 AGTGTACCATGTTATGAGACAGAGTTATCGAGCACAATCTCCACCACCTGTGCAGAGAA 868
DB 842 AGTGTGTAAGGGAACAGAGTGAACCTATCGGGGAATGTGCTTAACGGTGTCCGGGC 901
QY 869 GAACCTGCCAAGCTTGTATCTATGACACCACTGCATATGTCGAGCCCGAGAACTCT 928
DB 902 ACACCTGTACAGCTGAGTGCACAGACCCCTCACACATTAACAGAGACACCGAATAACT 961
QY 929 ACCCAATGCTGCTTGATCATGAACTACTGCAGAAATCCAGATGCTGTGGACCTCTT 988
DB 962 TCCCTGCAAAAATTGTGATGAATACTGTGCGCAATCTGACGAGAAAAAGGCCCAT 1021
QY 989 ATGTGTATACGAGGATCCCGGTGTACGTGGAGTACTGTCAACTGACGCAATGCTCAG 1048
DB 1022 GGTGCATTAACCAACCAAGCAGTGGGTGAGTACTGTAAAGATACCGT----- 1072
QY 1049 ACAGAGAAAGGACTGCGCTCGGCTCCGACCTGTACCCCGTTCAGACCTAGAGGCTC 1108
DB 1073 -----CTGTGACTCTCTCCCAATATCCAGC---GAACAAT 1105
QY 1109 CTTCGGAACAGACGACCTGAGCAAGAGCCGGGTGCAGAGATGCTACATGTTAATG 1168
DB 1106 TGGCTCCACAGCACCACCTGAGCTAACCCTGTGTGTCCAGAGACTGTAACATGTTGATG 1165
QY 1169 GACAGAGTTATGAGGACACATCTCCACACTGTGCAAGAAAGAACTGCCAAGCTTGT 1228
DB 1166 GACAGAGTTATGAGGACACATCTCCACACTGTGCAAGAAAGAAAGTGTGAGTCTGTGT 1225
QY 1229 CATCTATGACACGACCTGCAATGTCGAGCCCGCAAGATTAATCCCAATGCTGTGTA 1288
DB 1226 CATCTATGACACGACCGGACCAAGAGCCCAAGAAATCAACCAATGCTGTGCTCA 1285
QY 1289 TCATGAATCTGACAGAAATCCAGATGCTGTGGACCTCTTATTTATTCAGAGGATC 1348
DB 1286 CAATGAATCTGACAGAAATCCAGATGCTGTGGACCTCTTATTTATTCAGAGGATC 1345
QY 1349 CCGGTGTACAGTGGAGTACTGCAACCTGACAGCAATGCTCAGAGCAGAGAGGACTGCCG 1408
DB 1346 CCAGGCTCAGGTGAGTACTGTCAACCTGAAATAATGCTCAAGAAAGAAAGCGGTGTG 1405
QY 1409 TCAGGCTCAGGCTGTATACCCCGGTTCCAAAGCTTAAGAGCTCTTTCGAAACAGACCGCA 1468

```
Db 1406 TACACCTCCGCGCTGTTCCTGCTTCAGATGTAGAGACTCCCTCCGAAGAA----- 1458
QY 1469 CTGAGCAAAAGGCGCTGGGGTGCAGAGTGCTACCATGTGTAATGCACAGAGTTATTCAGGCA 1528
Db 1459 -----GACTGTATGTTGGAAATGGAAAGGATACCGAGCA 1495
QY 1529 CATACTCCACCATCTGCACAGGAAGAACCTGCAAGCTTGTCATCTATGCACACACACT 1588
Db 1496 AGAGGGGACACACTGTTCAGTGGAGCGCATGCCAGACTGGGGTGCACAGAGAGAGAGGCA 1555
QY 1589 CCGATAG---TCGGACCCCAATACATCTACCAATGCTGCTGATCATCACTACTGCA 1645
Db 1556 GACACAGCATTTTCACTCCAGAGCAAAATCCAGCGGGGCTGCGAAAAAATTTACTTGCC 1615
QY 1646 GGAATCCAGA---TGCTGTGCGACCTCCTTATGTATACAGAGGATCCCGGTGCAGGT 1702
Db 1616 GTACCCGTGATGTGATGTAGTGTGCTGCTGCTACAGACAAATCCAGAAACTTT 1675
QY 1703 GGGAGTACTGCACACTGCACGAATGCTCAGACGC 1736
Db 1676 ACGACTACTGTGATGTCCCTCAGTGTGGGCC 1709
```

RESULT 7

```
US-07-854-603-1
; Sequence 1, Application US/07854603
; Patent No. 5637492
```

GENERAL INFORMATION:

```
APPLICANT: Dawson, Keith M
APPLICANT: Edwards, Richard M
APPLICANT: Forman, Joan M
TITLE OF INVENTION: Activatable fibrinolytic and
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESSES:
ADDRESSER: Dr. John J. McDonnell
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
```

COMPUTER READABLE FORM:

```
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,603
FILING DATE: 19901207
```

CLASSIFICATION:

```
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,338
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
```

INFORMATION FOR SEQ ID NO: 1:

```
SEQUENCE CHARACTERISTICS:
LENGTH: 2753 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
```

FEATURE:

```
NAME/KEY: misc_feature
LOCATION: 1..2753
OTHER INFORMATION: /note="Fig. 2 Plasmidogen cDNA
FEATURE:
NAME/KEY: sig_peptide
```

```
LOCATION: 65..121
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 122..2494
FEATURE:
NAME/KEY: CDS
LOCATION: 65..2494
FEATURE:
NAME/KEY: misc_feature
LOCATION: 54..55
OTHER INFORMATION: /note="Ball site"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2565
OTHER INFORMATION: /note="Spli site"
US-07-854-603-1
```

Query Match

Best Local Similarity 4.7%; Score 236.6; DB 1; Length 2753;

Matches 410; Conservative 0; Mismatches 179; Indels 36; Gaps 3;

```
QY 92 CACGACGACCTGAGCAAGCATGTGTGTCAGAGATTGCTACCATGTGATGACAGAGTT 151
Db 1158 CAGACACACCTGAGTAACTCCCTGTGTGTCAGAGACTGTACATGTGATGACAGAGTT 1217
QY 152 ATGAGGCGACGTCTCCACCATCTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 211
Db 1218 ACCGAGGCGACGTCTCCACCATCTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1277
QY 212 CACCATCATACATANTAGAGACACAGAAACTACCCAAATGCTGCTGATGATGAGT 271
Db 1278 CACCATCATACATANTAGAGACACAGAAACTACCCAAATGCTGCTGATGATGAGT 1337
QY 272 ACTGAGGAGATCCAGATGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 331
Db 1338 ACTGAGGAGATCCAGATGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1397
QY 332 GGTGGAGTACTGCAACCTGACGCAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 391
Db 1398 GGTGGAGTACTGCAACCTGACGCAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1457
QY 392 CGACTGTACCCCGGTCCCAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 451
Db 1458 CGCTGTGTGCTGCTCCAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501
QY 452 GGCCTGGGGTGCAGAGTGCTACATGTGTAATGAGAGAGAGAGAGAGAGAGAGAGAG 511
Db 1502 -----GACTGTATGTTGGAAATGGAAAGGATACCGAGGAGAGAGAGAGAG 1547
QY 512 CCACGTGCACAGAGAAAGCTGCAAGCTGCTGATGATGATGATGATGATGATGATGAT 569
Db 1548 CCACGTGCACAGAGAAAGCTGCAAGCTGCTGATGATGATGATGATGATGATGATGAT 1607
QY 570 -TCGACCCCAAGTAATACCAATGCTGCTGATGATGATGATGATGATGATGATGAT 628
Db 1608 TTTTCATCCCAAGTAATACCAATGCTGCTGATGATGATGATGATGATGATGATGAT 1667
QY 629 A---TCGCTGTCAGCTCCTTATTTTATACAGAGAGATCCCGGTGATGAGTACT 685
Db 1668 ATGCTATGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1727
QY 686 GCAACCTGACGCAATGCTCAGACGC 710
Db 1728 GTGATGTCCCTCAGTGTGGGCC 1752
```

RESULT 8

```
US-07-750-080A-18
; Sequence 18, Application US/07750080A
; Patent No. 5445953
```

GENERAL INFORMATION:

```
APPLICANT: DORNER, F.
APPLICANT: SCHIEFLINGER, F.
```


REFERENCE/DOCKET NUMBER: 30472/166/1MMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2296 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid;
DESCRIPTION: Synthetic DNA oligonucleotide
IMMEDIATE SOURCE:
CLONE: pN29pt-LPg
US-08-651-472-18

Query Match 4.7%: Score 235; DB 3; Length 2296;
Best Local Similarity 58.5%; Pred. No. 4.2e-70;

Matches 546; Conservative 0; Mismatches 316; Indels 72; Gaps 5;

QY 467 AGTGTACCATGTAATGAGACAGATTATGAGGACATATCTCACAGGAA 526
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 610 AGTGTCTGAGGAGACAGTGAATACTGCGGGATGTGCTTTACGTTCCGGC 669
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 527 GAACGCGCAAGCTGTGTCATCTATGACACACACCTGCAATGCGACCCGAACTACT 586
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 670 AACCTGTGACAGCTGAGAGTGACAGACCCCTCACACATTAACAGSACACGAAACT 729
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 587 ACCCAATGCTGCTTGTATCATGAATCTGACAGAAATCCATGCTGTGGACCTCTT 646
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 730 TCCCTGCAAAAATTTGGATGATAAACTATGCTCCGCAATCTGACGAAAGAGCCCAT 789
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 647 ATTGTATGAGAGATCCGCTGTCAGGTGAGGTGAGTCTGACATGCAATGCTGAG 706
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 790 GGTGCATATCAACAAGCAAGTGGGTGGAGTCTGTAGATACCGT----- 840
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 707 AGCGAAGAGGACTCGCTCGCTCGACTGTACCCCGTTCCAGCTTAGAGGCTC 766
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 841 -----CCTGTGACTCTCTCCCAATGTCACG---GAACAAAT 873
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 767 CTTCCGAAACAGACCGATGAGCAAAAGGCTGGGGTGCAGAGTCTACATGTATAG 826
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 874 TGGCTCCACAGCACACCTGAGCTAACCCCTGTGTCAGAGCTGTACACAGGATG 933
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 827 GACAGAGTATGAGGACATATCTGACACACTGTCAACAGAAACCTGCCAAGCTTGT 886
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 934 GACAGAGCTACCGAGGACATCTCTCACACACACAGAAAGAGTGTCTGTGT 993
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 887 CATCTATGACACACACTCGATATGTGAGACCCCAATATCTACCAATGCTGCTTGA 946
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 994 CATCTATGACACACACCGGACCAAGAACCCCAAACTACCAAAATGCTGGCTGA 1053
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 947 TCATGAACTATGAGGAATCCAGATGTGTGAGAGCTCTTATTTATGAGAGGATG 1006
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1054 CATGAACTATGAGGAATCCAGATGTGTGAGAGCTCTTATTTATGAGAGGATG 1113
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1007 CCGGTGTCAGTGGGAGTACTGCAACCTGACAGCAATCTCAGACGAGAGGACTGCCG 1066
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1114 CCAGCTCAGTGGGAGTACTGCAACCTGCAAAATGCTCAGAGACAGAGGAGTGTG 1173
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1067 TGGCGCTCCGAGCTTTCACCGGTTCCAGGCTAGAGGCTCTCTCCGAAACGCCCA 1126
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1174 TACGACTCTCCGCTGTCTCTGCTCCAGATGTAGAGACTCTCTCCGAAAGAA----- 1226
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1127 CTGAGCAAAAGGCTGGGTGAGAGATGTACATGTGTAATGACAGAGTTATCGAGGCA 1186
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1227 -----GACTGTATGTTTGGGAATGGAAAGGATACCGAGCA 1263
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1187 CATACTCCACACTGTCTCAGAGAAAGCCCAAGCTTGGTCTATGACACCACTACT 1246
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1264 AGAGGGCGACACTGTCTTACTGAGAGCCATGCCAGAGTGTGCTGCCAGAGCCCATTA 1323
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111

QY 1247 CGCATAG---TCGACCCCAAGAACTATACCAAAATGCTGCTGTGATCATGAATCTGCA 1303
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1324 GACACGACATTTTACATCCAGACAAATCCACGGGGGTCTGAAAAAATTTACTGCC 1383
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1304 GGAATCCAGA--TGCTGTGACAGCTCTTATTGTTATACAGAGGATCCCGGTGCAGGT 1360
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1384 GTAACCTGATGATGATAGTGTGTCCTCGGTGTACACGACAAATCCAGAAATTTT 1443
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1361 GGGAGTACGCACTGACGCAATGCTCAGACCG 1394
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 1444 ACGACTACTGTGATGTCTCCCTCAGTGTGCGGCC 1477
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111

RESULT 10

US-08-358-928-18
Sequence 18, Application US/08358928
Patent No. 6265183

GENERAL INFORMATION:

APPLICANT: DORNER, Friedrich
APPLICANT: SCHEIFLINGER, Friedrich
APPLICANT: FALKNER, Falko Gunter
APPLICANT: PLEIDERER, Michael
TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF CHIMERIC
TITLE OF INVENTION: VIRUSES CONTAINING HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
TITLE OF INVENTION: (HIV-1) ANTIGENS
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/358,928

FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,738
FILING DATE: 20-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/750,080
FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/166/1MMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:
LENGTH: 2296 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid;
DESCRIPTION: Synthetic DNA oligonucleotide
IMMEDIATE SOURCE:
CLONE: pN29pt-LPg
US-08-358-928-18

Query Match 4.7%: Score 235; DB 4; Length 2296;

Best Local Similarity 58.5%; Pred. No. 4.2e-70;
Matches 546; Conservative 0; Mismatches 316; Indels 72; Gaps 5;

QY 467 AGTGTACCATGTAATGAGACAGATTATGAGGACATATCTCACAGGAA 526

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2542 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-441-370-1

Query Match

1.8%; Score 92.4; DB 1; Length 2542;
Best Local Similarity 98.9%; Pred. No. 2.2e-21;
Matches 93; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGGGATTGGACACACTTTCGACACTGCGCCAGTCCCAAAATGAGACATTAAGGAA 60
Db 1398 CTGGGATTGGACACACTTTCGAGCTGCGCCAGTCCCAAAATGAGACATTAAGGAA 1457

Qy 61 GTGGTCTCTACTCTTTTATTTCTGAATCAG 94
Db 1458 GTGGTCTCTACTCTTTTATTTCTGAATCAG 1491

RESULT 13

US-08-985-526-22
Sequence 22, Application US/08985526
Patent No. 6080728

GENERAL INFORMATION:

APPLICANT: Mixson, James A
TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
TITLE OF INVENTION: THERAPY
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Connolly, Bove, Lodge, & Hutz
STREET: 1220 Market Street, P.O. Box 2207
CITY: Wilmington
STATE: Delaware
COUNTRY: U.S.A.
ZIP: 19899

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,526
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,845
FILING DATE: 16-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: McMorrow Jr., Robert G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 658-9141
TELEFAX: (302) 658-5613
INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:
LENGTH: 645 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-985-526-22

Query Match

1.4%; Score 67.6; DB 3; Length 645;
Best Local Similarity 55.0%; Pred. No. 3e-13;
Matches 133; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 464 AGGAGTCTACATGTAATGACAGACTTATCGAGGACATCTCCACCACCTGTGCAG 523

Db 266 AGGATCATGATCTCTGACATGAGAAAGTATGAGGCAAAATCTCCAGACATCTGTG 325

Qy 524 GAGAACCTGCGACAGTTGGTATCATATGACACACACTGCGATAGTGGACCCAGAAAT 583

Db 326 GACTTGACTGCCAGGCTGGATTCAGAGCCACATGCTCAATGATCATCTCTGCCA 385

Qy 584 ACTACCCAAATGCTGCTTGCATGAACTACTGACAGAAATCCAGATGCTGTGCAGCTC 643

Db 386 AATTTCCACAGAACCTGAAGATGATATGTCACCAACCTGACGGGAGCCAGGC 445

Qy 644 CTATGTATGATGAGGATCCGCTGTCAGGTGAGTACCTGACACCTGACGCAATGCT 703

Db 446 CCTGCTCTTCACAAAGACACCCACCAACGCTGGAAATCTGTGACATCCCGCTGCA 505

Qy 704 CA 705

Db 506 CA 507

RESULT 14

US-08-985-526-24
Sequence 24, Application US/08985526
Patent No. 6080728

GENERAL INFORMATION:

APPLICANT: Mixson, James A
TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
TITLE OF INVENTION: THERAPY
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Connolly, Bove, Lodge, & Hutz
STREET: 1220 Market Street, P.O. Box 2207
CITY: Wilmington
STATE: Delaware
COUNTRY: U.S.A.
ZIP: 19899

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,526
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,845
FILING DATE: 16-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: McMorrow Jr., Robert G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 658-9141
TELEFAX: (302) 658-5613
INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:
LENGTH: 1284 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-985-526-24

Query Match

1.4%; Score 67.6; DB 3; Length 1284;
Best Local Similarity 55.0%; Pred. No. 4.4e-13;
Matches 133; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 806 AGGAGTCTACATGTAATGACAGACTTATCGAGGACATCTCCACCACCTGTGCAG 865

Db 905 AGGAATCAGTATCTGACAGTGAAGAAAGTATGAGGCAAAATCTCCAGACATGCTG 964

Qy 866 GAGAACCTGCGACAGTTGGTATCATATGACACACACTGCGATAGTGGACCCAGAAAT 925

Db 965 GACTTGACTGCCAGGCTGGATTCAGAGCCACCATGCTCATGATGATACATCCCTGCA 1024

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 06:40:43 ; Search time 758 Seconds
(without alignments)

4405.804 Million cell updates/sec

Title: US-09-923-515-3_COPY_1_5000
Perfect score: 5000
Sequence: 1 cttggatctggagacacattt.....caagcttgctcatatgac 5000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 478924 seqs, 333959956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1855.6	37.1	10422	9 US-09-870-759-125	Sequence 125, Appl
2	236.6	4.7	2497	10 US-09-946-893-1	Sequence 1, Appl
3	216.2	4.3	1134	10 US-09-873-676-29	Sequence 29, Appl
4	165.4	3.3	431	10 US-09-864-761-13941	Sequence 13941, A
5	145.2	2.9	422	10 US-09-833-381-252	Sequence 252, Appl
6	125	2.5	168	10 US-09-864-761-30505	Sequence 30505, A
7	118.4	2.4	160	10 US-09-864-761-20992	Sequence 20992, A
8	105.6	2.1	408	10 US-09-960-352-1535	Sequence 1535, Ap
9	93.2	1.9	1164	10 US-09-880-107-2422	Sequence 2422, Ap
10	90.2	1.8	242	10 US-09-960-352-13971	Sequence 13971, A
11	82.8	1.7	362	10 US-09-864-761-4339	Sequence 4339, A
12	67.6	1.4	645	12 US-10-036-869-22	Sequence 22, Appl
13	67.6	1.4	1284	12 US-10-036-869-24	Sequence 24, Appl
14	62.8	1.3	1344	9 US-10-081-309-1	Sequence 1, Appl
15	62.8	1.3	5898	10 US-09-880-107-3708	Sequence 3708, Ap
16	61.8	1.2	780	9 US-09-870-759-96	Sequence 96, Appl
17	58.8	1.2	402	10 US-09-960-352-15108	Sequence 15108, A
18	58	1.2	245	10 US-09-960-352-4911	Sequence 4911, Ap
19	57.6	1.2	214	10 US-09-960-352-1590	Sequence 1590, Ap

20	39.8	0.8	333	10 US-09-960-352-12502	Sequence 12502, A
21	38.2	0.8	176	10 US-09-864-761-31962	Sequence 31962, A
22	38.2	0.8	453	10 US-09-864-761-15448	Sequence 15448, A
23	36.8	0.7	216	10 US-09-960-352-12269	Sequence 12269, A
24	35.6	0.7	9980	10 US-09-880-107-3352	Sequence 3352, Ap
25	34.4	0.7	417	10 US-09-960-352-8663	Sequence 8663, Ap
26	33.6	0.7	451	10 US-09-964-824A-112	Sequence 112, Ap
27	33.6	0.7	10422	9 US-09-870-759-125	Sequence 125, Appl
28	33.2	0.7	11990	10 US-09-969-708-559	Sequence 559, Appl
29	33	0.7	440	10 US-09-895-828-293	Sequence 293, Appl
30	32.8	0.7	529	10 US-09-998-598-2598	Sequence 2598, Ap
31	32.8	0.7	556	10 US-09-919-580-316	Sequence 316, Appl
32	32.8	0.7	561	10 US-09-919-580-556	Sequence 556, Appl
33	32.8	0.7	564	10 US-09-919-580-310	Sequence 310, Appl
34	32.8	0.7	752	10 US-09-919-580-839	Sequence 839, Appl
35	32.8	0.7	756	10 US-09-919-580-821	Sequence 821, Appl
36	32.8	0.7	873	10 US-09-880-107-2349	Sequence 2349, Ap
37	32.8	0.7	1000	10 US-09-925-301-180	Sequence 180, Appl
38	32.6	0.7	372	10 US-09-960-352-7521	Sequence 7521, Ap
39	32.6	0.7	401	10 US-09-864-761-14212	Sequence 14212, A
40	32.6	0.7	8034	10 US-09-764-860-890	Sequence 890, Appl
41	32.4	0.6	418	10 US-09-983-965-713	Sequence 713, Appl
42	32.2	0.6	234	10 US-09-960-352-5589	Sequence 5589, Ap
43	32.2	0.6	389	10 US-09-960-352-6628	Sequence 6628, Ap
44	32.2	0.6	411	10 US-09-960-352-9849	Sequence 9849, Ap
45	32	0.6	2000	9 US-09-938-842A-2905	Sequence 2905, Ap

ALIGNMENTS

RESULT 1
US-09-870-759-125
Sequence 125, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: TERMAN, David S
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT APPLICATION NUMBER: US/09/870,759
CURRENT FILING DATE: 2002-01-14
PRIOR APPLICATION NUMBER: US 60/208,128
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PatentIn version 3.1
SEQ ID NO 125
LENGTH: 10422
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (6667)..(10176)
OTHER INFORMATION:
US-09-870-759-125

Query Match	37.1%	Score 1855.6;	DB 9;	Length 10422;
Best Local Similarity	83.0%	Pred. No. 0;		
Matches 2138; Conservative	0;	Mismatches 414;	Indels 24;	Gaps 1;
OY 1623 TGGCTTGATCTGAACTACTGACAGATCCAGATGCTGGAGCTCCCTATTGTTATAC	1682			
Db 6657 TGGCTTGATCTGAACTACTGACAGATCCAGATGCTGGAGCTCCCTATTGTTATAC	6716			
OY 1683 GAGGATCCCGGTGTCAGGTGGAGTACTGCAACTGACCAATGCTCAGACGACAGAG	1742			
Db 6717 GAGGATCCCGGTGTCAGGTGGAGTACTGCAACTGACCAATGCTCAGACGACAGAG	6776			
OY 1743 GACTCCGTCGCGCTCGACTGTTACCCGGTTCGAAGCTGAGGCTCCCTCCGACAA	1802			
Db 6777 GACTCCGTCGCGCTCGACTGTTACCCGGTTCGAAGCTGAGGCTCCCTCCGACAA	6836			
OY 1803 AGCACCGACTGAGCAAAAGCGCTGGGGTGCAGAGATGCTTACCATGTGAATGACAGATTA	1862			

Db 6837 AGCAACAACGTGACAAAGGCGTGGGGTGCAGGAGTCTTACCACGAAATGACAGATTGA 6896
QY 1863 TCGAGGACATACTCCACACACTGTGCACAGAGAAGAACCTGGCAAGCTTGTGATCTATGAC 1922
Db 6897 TCAAGGACATACTCTATTACTGTGCACAGAGAAGAACCTGGCAAGCTTGTGATCTATGAC 6956
QY 1923 ACCCACTCCGATGTGGGAGCCCGAGAAATCTACCCAAATGCTGGCTGTGATCTGAACCTA 1982
Db 6957 ACCCACTCCGATGTGGGAGCCCGAGAAATCTACCCAAATGCTGGCTGTGATCTGAACCTA 7016
QY 1983 CTGCAAGAAATCCAGATCTGTGAGAGCTCTTATTGTTATACGAGGATCCCGGTGTACG 2042
Db 7017 CTGCGAAATCCAGATCTGTGAGAGCCCTTGGTTATACCAAGATCCCGGTGTACG 7076
QY 2043 GTGGGATGATGCAACTGTACGCAATGCTACAGCGCAAGAGGATGCGGCTGCGCCTCC 2102
Db 7077 GTGGGATGATGCAACTGTACGCAATGCTACAGCGCAAGAGGATGCGGCTCTCCCTCC 7136
QY 2103 GACTGTTACCCCGGTTCCAGAGCTCCAGAGGCTCCTTCGCAACGACCGACTGAGCAAG 2162
Db 7137 GAATGTTATCTGTGCTCCAGAGCTTGAAGGCTTTTGTGAACAGCACTGACTGAGCAAG 7196
QY 2163 GCGTGGGCTGACAGAGTGTACCATGTTAATGACAGAGTTAATCGAGGACATATCTCAC 2222
Db 7197 CCGCGGGGTACAGGACTGTACTACCATTTATGACAGAGTTACCGAGGACATATCTCAC 7256
QY 2223 CACTGTACAGGAACCTGGCAAGCTGTGATCATCTATGACACGACCTCGCATAGTGC 2282
Db 7257 CACTGTACAGGAACCTGGCAAGCTGTGATCATCTATGACACGACCTCGCATAGTGC 7316
QY 2283 GACCCCAAGAAATCTACCCAAATGCTGGCTTATGATGAACTACTGCAAGAAATCTGCA 2342
Db 7317 GACCCCAAGAAATCTACCCAAATGCTGGCTTATGATGAACTACTGCAAGAAATCTGCA 7376
QY 2343 TGTGGCAGCTCTCTATTGTTATGATGAGGAGTCCGGGTGACAGTGGGAGTACGCAACT 2402
Db 7377 TGTGATTCGGCCCTTGGTGTACACATGATCCCAAGTGTGAGTGGAGTACGCAACT 7436
QY 2403 GACCAATGCTCAGACGACAGAGGAGTCCGCTGCGGCTCCGACTGTGAACCCCGGTTCC 2462
Db 7437 GACCAATGCTCAGTGTGACAAATCAAGTCTCTGCAACTCTCAGGTGTCCCAATGTC 7496
QY 2463 AAGCCTAGAGGCTCTTCGCAACAGACGACGACTGAGCAAGAGCCGCGGGTGTGAGAGTC 2522
Db 7497 AAGCAGAGAGGCTTCTTCTGAGAGAGACCAAGAGCAAGAGCCCGGGGTGTGAGAGTC 7556
QY 2523 CTACCAATGTTAATGACAGAGATTATGAGGACATCTCCACCACTGTGACAGAGAAAG 2582
Db 7557 CTACCAATGTTAATGACAGAGATTATGAGGACATCTCTACCACTGTGACAGAGAAAG 7616
QY 2583 CTGCGAAGCTTGTGCAATCTATGACACCACTGTGATAGTGGAGCCCGCAAGATCTACCC 2642
Db 7617 ATGTCATCTTGTGCTCTTATGACACCACTGTGATAGTGGAGCAACCAATATATATCTC 7676
QY 2643 AAATGCTGCTGATCATGAACTACTGAGGAATCCAGATGCTGTGGAGCTCTCTATTG 2702
Db 7677 AAATGCTGCTGATCATGAACTACTGAGGAATCCAGATGCTGTGGAGTATAGTCTTGGTG 7736
QY 2703 TTATACAGAGGATCCCGGTGTGAGTGGAGTACTGCAACTGTGAGCAATGCTCAGAGCC 2762
Db 7737 TTATACAGGATCCCAACTGATGAGTGGAGTACTGCAACTGTGAGCAATGCTCAGAGCC 7796
QY 2763 AGAAGGATGCGGCTGCGGCTCCGACTGTTAACCCGGTTCCAAAGCTGTAGAGGCTCTTC 2822
Db 7797 AGAATCAAGTGTCTCTGCGAGCTCAC-----GGCTGTTTC 7832
QY 2823 CGAACAAGACCGACTGAGCAAAAGGCTGGGTGACAGAGTGTACCATGTTAATGAGCA 2882
Db 7833 TGAACAAGACCGAAGAGAGCCCAAGTCCAGAGCTGTACCATGTTAATGAGAGCA 7892
QY 2883 GAGTATACGAGGACATATCTCCACCTGTGACAGAGAAAGACATGTCAAGTCTTGGTCTC 2942
Db 7893 GAGTATACGAGGCTCAATCTCCACCTGTGACAGAGAAAGACATGTCAAGTCTTGGTCTC 7952

QY 2943 TATGACACCAACTGCGATAGTCCGACCAGAAATACTACCCAAATGCTGGCTTATCAT 3002
Db 7953 TATGACACCAACTGCGATAGTCCGACCAGAAATACTACCCAAATGCTGGCTTATCAT 8012
QY 3003 GAATCTACTGCAAGAAATCCAGATGCTGTGGAGCTCTTATTGTTATACAGAGGATCCCGG 3062
Db 8013 GAATCTACTGCAAGAAATCCAGATGCTGTGGAGCTCTTATTGTTATACAGAGGATCCCGG 8072
QY 3063 TGTGAGTGGAGTACTGCAACTGAGCAATGCTCAGAGCAAGAGGAGCTGCGCTGCG 3122
Db 8073 TGTGAGTGGAGTACTGCAACTGAGCAATGCTCAGAGCAATGCTGATGSAATCAACTCTCTCCAC 8132
QY 3123 GCGTCCGACTGTACCCCGGTTCCAAAGCTTGTGAGAGCTCTTCCGAAACACACCTGCA 3182
Db 8133 AACTCCACAGGTTGGTCCAGGTTCCAAAGCAAGAGCTTCTCTGTAAGAACACCACTGTA 8192
QY 3183 GCAAAGGCTGGGGTGCAGAGTGTACCAATGTTAATGACAGAGTTATGAGGACATTA 3242
Db 8193 AAGCAGCACTGGGGTCCAGAGCTGTACCGAGTGTATGACAGAGTTATGAGGACACT 8252
QY 3243 CTCCACCACTGTACAGAGAAACCTGCCAAGCTTGTGATCTATGACACCACTGCA 3302
Db 8253 CTCCACCACTGTACAGAGAAACATGTGATGCTTGGTGTATGACACCACTGCA 8312
QY 3303 TAGTGGAGCCCAAGATCTACCAATGCTGGCTGTATGATGAACCTGACAGAGAAATCC 3362
Db 8313 TCGAGAGATCCCAATTAATCTATCCAAATGCTGGCTGACAGAGAACTGTGAGAGATTC 8372
QY 3363 AGATGCTGTGACGCTCTCTTATTGTTATACAGAGATCCCGGTGTGAGTGGAGTACTG 3422
Db 8373 AGATGCTGTGAGTGGCTCTTGTGTATACCATGATGCCAGTGTGAGTGGAGTACTG 8432
QY 3423 CAACCTGACGAATGTCTCAGAGCAAGAGGAGCTGGCGCTCCGACTGTACCC 3482
Db 8433 CAACCTGACGAATGTCTCAGAGCAAGAGTGTCTCAGAACTCCCACTGCGCC 8492
QY 3483 GGTTCGAAGCTTAGAGGCTCTTCGGAACAAGCAACCACTGAGCAAAAGGCTGGGTGCA 3542
Db 8493 GGTTCGAAGCAAGAGGCTCTCTCTGAAACAAGCAACCACTGAGAAAGGCTGGGTGCA 8552
QY 3543 GGAAGTACCAATGTTAATGACAGAGTATCCAGAGGACATCTCCACCACTGTCAAG 3602
Db 8553 GGAATGTACCAATGTTAATGACAGAGTATCCAGAGGATATCTCCACCACTGTCAAG 8612
QY 3603 AAGAACTGCGCAAGCTTGTATCTATGACACCACTGCAATGTTGAGAGCCCAAGATA 3662
Db 8613 AAGGACCTGTCAATCTTGTGTATCTATGATACCACTGCAATGAGAGCCCAAGATA 8672
QY 3663 CTACCCAAATGCTGGCTGTGAGAGTACTGTGAGAAATCCAGATCTGGCAAAAC 8732
Db 8673 CTACCCAAATGCTGGCTGTGAGAGTACTGTGAGAAATCCAGATCTGGCAAAAC 8732
QY 3723 TTATGTTATACAGAGGATCCCGGTGTGAGTGGAGTACTGCAACTGCAAGATGCTC 3782
Db 8733 CTGGGTGTACCAACCACTGCTGTGTGTGAGTGGAGTACTGCAATGTGACAAATGCTC 8792
QY 3783 AGACGAGAAAGGAGTCCGCTGCGCTCCGACTGTTACCCGGTTCCAAAGCTTAGAGC 3842
Db 8793 AGAAACAGAAATCAAGTGTCTGAGAGTCTCCCACTGTTGTTCCAGTTCCAAAGCATGAGAGC 8852
QY 3843 TCCCTCCGAACCAACGACGCACTGAGCAAGGCTCCGGGGTCCAGAGGCTACCATGTTA 3902
Db 8853 TCATTTGTGAGACACCACTGAGCAAAACCTGTGTGTGCGGAGTGTACCATGTTA 8912
QY 3903 TGCACAGAGTTATGAGGACATCTCCACACTGTGTACAGAGAAAGCACTGCAAGCTTG 3962
Db 8913 TGCACAGAGTTATGAGGACATCTCCACACTGTGTACAGAGAAAGCACTGCAATCTTG 8972
QY 3963 GTCACTTATGACACCACTGCACTGTGTGAGGACCCCAAGATACTACCCAAATGCTGCTT 4022
Db 8973 GTCACTTATGACACCACTGCACTGTGTGAGGACCCCAAGAACTACCCAAATGATGCTCT 9032

```

OY 4023 GATATGAACACACGACGAAATCCAGATGCTGAGCAGCTCCCTATTTGTTATACGAGGA 4082
|| |||||
DB 9033 GACATGAACACTACTGCAGGAATCCAGATCCGATACAGGCCCTTGTTTACCATGGA 9092
|| |||||
OY 4083 TCCCGGTTCAGTGGGAGTAGTACTGCACACCGACGCAATGCTTCAGACGAGAAAGGACTGC 4142
||| || |||||
DB 9093 CCCGAGCATCAGGTGGGAGTAGTCTCAACCTGACGCGCATGCTTCAGACACGAAGAAGGACTGT 9152
|| |||||
OY 4143 CGTGGCGCTTCGACTGTTCACCCGGTTCCAGGCTTAGAGGCTTCCTTCGAGACAAG 4198
||||| |||||
DB 9153 GGTGCGCTTCGACTGTCACTCAGGTCCAGAGCTTCGGGCTTCCTTGAACAAG 9208
|||||

RESULT 2
US-09-946-893-1
; Sequence 1, Application US/09946893
; Patent No. US20020072494A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
; TITLE OF INVENTION: Inhibitors
; FILE REFERENCE: Newburn
; CURRENT APPLICATION NUMBER: US/09/946, 893
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230, 893
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(2482)
; US-09-946-893-1

```

Query Match	4.78;	Score 236.6;	DB 10;	Length 2497;
Best Local Similarity	58.68;	Pred. No. 3.3e-67;		
Matches 547; Conservative	0;	Mismatches 315;	Indels 72;	Gaps 5

OY	809	AGCTTACCGTGGTAATGGACAGAGTTATTCGAGGACACTACTCCACACCTGTCACAGAA	868
Db	870	AGTGTCTGAAAGGAAACAGGTGAAAACTATTCGGGGAAATGTGGCTGTAACTCGTGGGGC	929
OY	869	GAACCTGCACAGCTGTGGTCACTATGATACACACACACTCCGACTAGTCGAGACCCAGAAATCT	928
Db	930	ACACTCTTCAGCACTTGGATGGACAGACACCCCTTCACACACTATACAGSACACACGAAACACT	969
OY	929	ACCCAAATGCTGGCTTGATCATGAACTACTGTCAGAAATCCAGATGCTGTGGCAGCTCTTT	988
Db	990	TCCCTTCGAAAAATTTGGATGAAAAACTACTGTCCGCAATCTGACAGGAAAAAGGGCCCCAT	1049
OY	989	ATTGTATTAGAGAGATCCCGGTGTGAGTGGGAGTACTGTGCACCTGACGCAATAGCTCAG	1048
Db	1050	GGTGCCTAATACCAACCAACAGCCAAAGTGCSTGGGAGTACTGTAAAGTACCGT-----	1100
OY	1049	ACGCAAGAGGACTGCGCTGCGGCTCCGCACTTTACCCCGGTTCCAAAGCTTAGAGGCTC	1108
Db	1101	-----CCTGATCACTCTCCCACTATTCAGG-----GAACAT	1133
OY	1109	CTTCCGAAACAGACCCGACTGAGACMAAGGCCCTGGGGGTCCAGAGTGTACTACCAATGTAATG	1168
Db	1134	TGGCTCCACAGACCACTGAGACTTAAACCCCTGTGGTCTCCAGAGCTGCTACCACTGTGTATG	1193
OY	1169	GACAGATTTATCGAGGACATACTCCACCACTGTACACAGGAAGAACTGCCAAGCTGGT	1228
Db	1194	GACAGAGCTACCGGAGGACATCTCTCCACCAACCAACAGAAAGAAAGTGTGAGTCTTGGT	1253
OY	1229	CATCTATGACACACACTGCTGATATGTGCGAACCCAGAAATCTACTACCAATATGTGCTTGA	1288
Db	1254	CATCTATGACACACACCGGACACAGAAAGCCCGAAAAATCTACCAATATGTGCTTGA	1313

QY	1289	TCGGAAGCTACTGGAGGAATCCGAGAGCGCTGGCAGCTCCTATTGTTATACGAGGATC	1348
DB	1314	CAATGAAGTACTGCGAGGAATCCGATGACGATTAAGGCCCTGTGTGTTTACCGAGACC	1373
QY	1349	CCGGTGTAGAGTGGGAGTACTGCAACCTGACGCGATGCTCAGACCCAGAGGAGCTGCG	1408
DB	1374	CCAGGTGATGAGTGGGAGTACTGCAACCTGAAAAAATGCTCAGSAAACGAGGAGGTGTG	1433
QY	1409	TGCGGCTTCGACGTGTACCCCGGTTCCAAAGCTTAAGGCTCTTCCGACAGACCGGA	1468
DB	1434	TAGCACTCCGCCCTGTTGTCTCTGCTTCCAAATGTAAAGACTCTTCCGAAGAA-----	1486
QY	1469	CTGAGCAAAAGGCGTGGGGTGCAGAGAGTGCATCCATGATGTATGACAGATTATCAGGCA	1528
DB	1487	-----GACTGTAATGTTTGGGAATGGGAAGATATCCGAGCA	1523
QY	1529	CATTAATCCACCACTGTACAGAGAAAGCTGCCAGACTTGTATCTATGACACCAACT	1588
DB	1524	AGAGGGCGACCACTGTACTGTGGGACGCCAATCCGAGACGTGGCTGCCAGGAGCCCATTA	1583
QY	1589	CGCATG---TCGGAGCCCGAAGAAATCACTACCAATCTGCTGGTGTGATCATGAACTACGCA	1645
DB	1584	GACACAGCAATTTTCACTACCTCCAGACAAATCCAGGGGGGTCTGGAAAAAATTTACTGCC	1643
QY	1646	GGAATCCAGA---TGTGTGGCAGACTCCTATTGTTATACGAGGAGTCCCGGTGTCAAGT	1702
DB	1644	GTAACCTCGATGATGATGTAGGTGTCCCTGATGTACACGAAATCCAGAAAATCTTT	1703
QY	1703	GGGAGTACTGCACACTGACGCAATGCTCAAGCC	1736
DB	1704	ACGACTACTGATGTCCCTCAGGTCCGGCCCC	1737

```

?      RESULT 3
?      US-09-873-676-29
?      Sequence 29, Application US/09873676
?      Patent No US20020077289A1
?      GENERAL INFORMATION:
?      APPLICANT: Macdonald, Nicholas J.
?      APPLICANT: Sim, Kim L.
?      TITLE OF INVENTION: Angiotensin and Endostatin Binding Proteins and Methods of Use
?      FILE REFERENCE: 05213-0378 (43170-259333)
?      CURRENT APPLICATION NUMBER: US/09/873,676
?      CURRENT FILING DATE: 2001-06-04
?      PRIOR APPLICATION NUMBER: US 60/209,065
?      PRIOR FILING DATE: 2000-06-02
?      PRIOR APPLICATION NUMBER: US 60/289,387
?      PRIOR FILING DATE: 2001-05-08
?      NUMBER OF SEQ ID NOS: 123
?      SOFTWARE: PatentIn version 3.1
?      SEQ ID NO 29
?      LENGTH: 1134
?      TYPE: DNA
?      ORGANISM: Homo sapiens
?      US-09-873-676-29

```

Query Match	4.38;	Score 216.2;	DB 10;	Length 1134;
Best Local Similarity	80.18;	Pred. No. 1e-60;		
Matches 254; Conservative	0;	Mismatches 63;	Indels 0;	Gaps 0;

QY	92	CAGCAGCACTGACGAAAGCCATGTGGTCCAGATTGCTACCATGSGTATGACAGACTT	153
Db	818	CAGCACCACCTGAGCTAACCCCTGTGGTCCAGGACTGCTACCATGGTGATGACAGAGCT	877
QY	152	ATCAGGCGACGTACTCTCACCACGTGCACAGGAAGAGACTGCCAAGTGTGTCATCTATGA	211
Db	878	ACCGAGGCAATCTCTCCACACACACACAGGAAAGATGTCAGTGTGGTATCTATGA	937
QY	212	CACGACATCACTAATATAGGACACAGAAAATCTCCCAATGTGCTGATGATGAACCT	274
Db	938	CACACACCGGCGACCGAAGAGACCCCGAAGAACTACCCAAATGCTGGCTCGACATGAACCT	997
QY	272	ACTCGAAGAAATCCAGATGCTGTGGCAGCTCTCTTATTTGTATACGAGGATCCCGGTGTCA	331

Accession	Sequence	Position
Db	ACTCGAGGAATCCGAGATGCCGATTAAGGCCCCCTGGTGTTTACACAGACCCACAGCTTA	1057
QY	GGTGGGAGTACTGCAACCTTAGCAATGCTCAGACGCAAGAGGATGCGTGGCGCTC	391
Db	GGTGGGAGTACTGCAACCTTAAGAAAAATGCTCAGAGAACAGCATGTGTTAGCACCTC	1117
QY	CGACTGTTACCCGGTT	408
Db	CGCTGTGTGTCGTCTT	1134

RESULT 4
US-09-864-761-13941
; Sequence 13941, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION.

```

: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```

;
; CURRENT APPLICATION NUMBER: US/09/864,761
;
; CURRENT FILING DATE: 2001-05-23
;
; PRIOR APPLICATION NUMBER: US 60/180,312

?	PRIOR APPLICATION NUMBER: US 60/207,456	?	TITLE OF INVENTION: No. US20020132090A1
?	PRIOR FILING DATE: 2000-05-26	?	FILE REFERENCE: 5800-119
?	PRIOR APPLICATION NUMBER: US 09/632,366	?	CURRENT APPLICATION NUMBER: US/09/833,388
?	PRIOR FILING DATE: 2000-08-03	?	CURRENT FILING DATE: 2001-04-11
?	PRIOR APPLICATION NUMBER: GB 24263, 6	?	PRIOR APPLICATION NUMBER: 09/516,448
?	PRIOR FILING DATE: 2000-10-04	?	PRIOR FILING DATE: 2000-02-29
?	PRIOR APPLICATION NUMBER: US 60/236,350	?	

```

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 252
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-252

```

?	Query Match	2.9%	Score 145.2;	DB 10;	length 422;
?	Best Local Similarity	87.4%;	Pred. NO. 1.2e-37;		
?	Matches 159; Conservative	0;	Mismatches 23;	Indels 0;	Gaps 0;
?	Prior Application Number: PCT/US01/00665				
?	Prior Filing Date: 2001-01-30				
?	Prior Application Number: PCT/US01/00669				

PRIOR APPLICATION NUMBER: C/T/US01/00668	
PRIOR FILING DATE: 2001-01-30	
PRIOR APPLICATION NUMBER: PCT/US01/00663	
Qy	255
TGGCTTATATATACTACGCGAGGAATCCAGATGCTGGCAGCTCCTTATTGTATAC	31
Db	150
TGAATTATATACGATCGCAGGATCCAGATCCCTGGGAGGCGCTTGGGATTAAC	209

[illegible]

ENION FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670

? PRIOR FILING DATE: 2000-09-21
 ? PRIOR APPLICATION NUMBER: US 09/608,408
 ? PRIOR FILING DATE: 2000-06-30
 QY 435 Ag 436
 11

```

;
; PRIOR FILING DATE: 2001-01-29
;
; NUMBER OF SEQ ID NOS: 49117
;
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;
RESULT 6

```

```

; LENGTH: 431
; TYPE: DNA
; ORGANISM: Homo sapiens
US_09_004782.96205
; Sequence 30505, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION

```

1	FEATURE:	
2	OTHER INFORMATION: MAP TO AL109933.8	APPLICANT: Penn, Sharon G.
3	OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5	APPLICANT: Rank, David R.
4		APPLICANT: Hanzel, David K.

US-09-864-761-13941

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

	Query Match	3.3%;	Score 165.4;	DB 10;	Length 431;
	Best Local Similarity	89.4%;	Pred. No. 2.9e-44;		
	Matches 178;	Conservative	0;	Mismatches 21;	Indels 0;
				Gaps	
Qy	AGGAAAGGAGTCTTCTACTCTTTTATTTTGAATTCAGACAGCCTGAGCAAAAGCATG	11			
Db	232 AAGAACTATTCAGCTTGATTTCTTCTGTTTAAATTTAGACAGCACTGAGCAAAAGCATG	29			
Qy	116 TGGGCCAGAGATTGCTACCATGGTGTATGGACAGATTTTCGAGGACAGTCTCCACACTG	17			
Db	292 TGGTCCAGAGATTCTACCATGGTGTATGGACAGATTTTCGAGGACAGTCTCCACACTG	35			
Qy	176 TCACAGGAAGAGACAGCCAGCTTGCTATCTATGACACACATCTCAATATATAGACCA	23			
Db	352 TCACAGGAAGAGACCTGCAAGCTTGCTATCTATGACACACATCTCAATATATAGACCA	41			
Qy	236 CAGAAAACATCCCAATATGC	254			
Db	412 CAGAAAACATCCCAATATGC	430			

RESULT 5
US-09-833-381-252

; Patent No. US20020132090A1
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Robison, Keith E.

```

; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11

```

```

;
; INBOX REPLICATION NUMBER: 0/210/440
;
; PRIOR FILING DATE: 2000-02-29
;
; NUMBER OF SEQ ID NOS: 2050
;
; SOFTWARE: FastSeq for Windows Version 3.0

```

```

/ SEQ ID NO 252
/
/ LENGTH: 422
/
/ TYPE: DNA
/
/ COMMENT: "not available"

```

US-09-833-381-252

Query Match 2.98; Score 145.2; DB 10; Length 422;

Matches 159; Conservative 0; Mismatches 23; Indels 0; Gaps 0

255 TGGCTTCATCATGAACCTACTGCAGGAATCCAGATGCTGTGGCAGCTCCCTTATTGTTATAC 314

Db	150	TCACCTTCATCATGA	CTACTCGAGGAATC	GATCCTGTGGAAGG	CCCTTGGTGTTACAC	209
Ov	315	GAGGGATCCCGG	GTTCAGGTTGGGAGT	ACTGCCAACCTGAC	CGCAATGCTCAGAC	374

Db 210 GATGGATCCCAAGTCAGATGGAGTACTGCACCTGACACGATGCTCAGACACAGAAG 26

Ov 375 CAACTCGCCGCCCGCCGCGCCGCACTCCCTTAACCGCCGCTTGCAACCGTACAGCGGCGCTTCCGCAACA 43

Db 270 GACTGACGTCGTGCTCTGACTGTTATCCCGGTTCACAGCCTAGAGGATCCTTCCAAACC 322

Db 330 Ag 331

RESULT 6
US-09-864-761-30505
Sequence 30505. Annotation MS/09864761

```

; PATENT NO.0320020046703A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Penn, Sharron G.
; INVENTOR: Penn, Sharron G.

```

; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR


```
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30505
; LENGTH: 168
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109933.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: NT HIT: M86879.1, EVALUATE 2.00e-82
; OTHER INFORMATION: EST_HUMAN HIT: R07514.1, EVALUATE 2.00e-30
; OTHER INFORMATION: SWISSPROT HIT: P08519, EVALUATE 8.00e-22
US-09-864-761-30505

Query Match          2.5%; Score 125; DB 10; Length 168;
Best Local Similarity 100.0%; Pred. No. 3.1e-31;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
US-09-864-761-20992/c
; Sequence 20992, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20992
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109933.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.8
; OTHER INFORMATION: EST_HUMAN HIT: T89505.1, EVALUATE 4.00e-85
; OTHER INFORMATION: SWISSPROT HIT: P00747, EVALUATE 2.00e-21
; OTHER INFORMATION: NT HIT: M74220.1, EVALUATE 3.00e-85
US-09-864-761-20992

Query Match          2.4%; Score 118.4; DB 10; Length 160;
```


Db	266	AGGAATGCAATGATCTAGCTGGAGAAAAGTATGAGGGCAAAATCTCCAAAGACATGTCTG	325
Qy	524	GAAGAACCCTGCCAACCTTGGTCATCTATGACACCACACATCGATAGTGGACCCCAAAAT	563
Db	326	GAACTTGACCTGCGCAGGCGCTGGGATTTCTCAGAGCCCAACATCTCATGTGATACATCCCTGCCA	385
Qy	584	ACTACCCAAATGCTGGCTTGATCATGACATGACACTGCGAGAAATCCAAATCCTGTGGCAGTC	643
Db	386	AATTTCCAAAGCAAGAACTGAAAGATGAAATTAATTGCCAACCCCTGACGGGGAGCCAAAGGC	445
Qy	644	CTTATTTGTTATATCGAGGAGTCCCGGTCTCAGGTGGGAGTACTGCAACTGACGCATAGCT	703
Db	446	CCCTGGGTCTTCAACAACAGACCCACCAAAAGCTGGGAATATGTGATCATCCCCCGCTGCA	505
Qy	704	CA 705	
Db	506	CA 507	

RESULT 13
 US-10-036-869-24
 : Sequence 24, Application US/10036869
 : Patent No. US20020151516A1
 GENERAL INFORMATION:
 APPLICANT: Mixson, James A
 TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
 ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
 THERAPY
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Connolly, Bove, Lodge, & Hutz
 STREET: 1220 Market Street, P.O. Box 2207
 City: Wilmington
 STATE: Delaware
 COUNTRY: U.S.A.
 ZIP: 19899
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/036,869
 FILING DATE: 29-NO. US20020151516A1-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/965,526
 FILING DATE: <Unknown>
 APPLICATION NUMBER: US 08/608,845
 FILING DATE: 16-JUL-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: McMorrow Jr., Robert G
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (302) 658-9141
 TELEFAX: (302) 658-5613
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1284 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 24:
 US-10-036-869-24

	Query Match	1.4%	Score 67.6	DB 12	length 1284
	Best Local Similarity	55.0%	Pred. No. 6.5e-12		
	Matches 133	Conservative	0	Mismatches 109	Indels 0
				Gaps	0
Qy	AGAGAGCTCTACCATGATGATGAGACAGAGTATATCAGAGACAT	ACTATCCACCATCTACAG	865		
Db	AGGATATCATGATCTGCTGCTGAGGAGAAAGTATGAGGAGAAATCT	CCAGACATCTGCTG	964		
Qy	866	GAGACACTTCCAGAGCTTGCTGATCTATATGACACACCACTGC	GCATATGTTGGAGCCCAAGAT	925	

[illegible]

```

US-10-081-309-1
US-10-081-309-1
Sequence 1, Application US/10081309
Publication No. US20030012775A1
GENERAL INFORMATION:
APPLICANT: Hoffmann-La Roche Inc.
TITLE OF INVENTION: PEG Conjugates of NK
FILE REFERENCE: 20859
CURRENT APPLICATION NUMBER: US/10/081,309
CURRENT FILING DATE: 2002-06-04
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1344
TYPE: DNA
ORGANISM: homo sapiens
US-10-081-309-1

```

Query Match	1.3%	Score 62.8	DB 9	Length 1344
Best Local Similarity	53.7%	Pred. No. 2.5e-10		
Matches 139	Conservative	0	Mismatches 112	Indels 0
			Gaps	0
QY	124	GATTGCTACCATGTGTATGACAGAGATTATGACAGGACAGCTGCTCCACCAGCTGTCACAGGA	183	
Db	817	GAATGGCTCCMAAGTTCAGAGAAAGGCTTCACAGGGACCTGTCATTAATTCATTGGAAATGA	876	
QY	184	AGAGACCTGCCACCTTGGTCATCTATGACACCACATCAACATTAATAGACACAGAAAC	243	
Db	877	ATTCCATGTCAGCGTTGGGATTCTCAATATCCTCACAGACATGACATGACTCCTGAACAAAT	936	
QY	244	TACCCAAATCTGCGTTGATTCATGAACTACACAGAGATCCAGATGCTGTGGACGCGCT	303	
Db	937	TTCAAGTGCAGAGACCTTCAGAGAAATTTCTGTCGAAATTCAGATGGGTCTGAATACCC	996	
QY	304	TATTTTATACGAGGAGATCCGGGTGTACAGTGGGAGTACTGACAACTGACAGCAATGCTCA	363	
Db	997	TGGTGTTCACACTGATCCAAACATCCGAGTGGTGGCTCTCCCAAAATTCCAAACCTG	1056	
QY	364	GA 365		
Db	1057	GA 1058		

RESULT 15
US-09-880-107-3708
Sequence 3708, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIORITY APPLICATION NUMBER: US 60/211,379

```
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3708
; LENGTH: 5898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X16323
US-09-880-107-3708
```

```
Query Match      1.3%; Score 62.8; DB 10; Length 5898;
Best Local Similarity 53.7%; Pred. No.5.7e-10;
Matches 130; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

```
QY 124 GATTGCTACCATGATGATGAGCAGAGTTATCGAGGACGCTACTCCACCACTGTACAGGA 183
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1044 GAATGCATCCAGSTCAAGSAGAAGGCTACAGGGGCACTGCCAATACATTTGGAAATGGA 1103

QY 184 AGGACCTGCCAAGCTTGCTCATCTATGACACCAATCAACATTAATAGACACAGAAAC 243
      - - - - - ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1104 ATTCCATGTCAGCGTGGGATTTCTCAGTATCTCACAAGCATGACATGACTCTGAAAT 1163

QY 244 TACCCAAATGCTGCTGATCATGAATCACTACTGAGAGATCCAGATGCTGTGGCAGCTCT 303
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1164 TTCAAGTGCAGAGACTACGAGAAATTAATCTGCCGAAATCCAGATGGGTCTGAATCACC 1223

QY 304 TATGTTATACGAGGATCCCGGTGTCAGGTGGGAGTACTGCAACCTGACGCAATGCTCA 363
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1224 TGGTGTTTACACTGATCCAAACATCCGAGTTGGTTACTGCTCCCAATTCCAAACCTGT 1283

QY 364 GA 365
      ||
Db 1284 GA 1285
```

Search completed: March 6, 2003, 15:50:52
Job time : 826 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 05:15:44 ; Search time 5999.5 Seconds
(without alignments)
13497.364 Million cell updates/sec

Title: US-09-923-515-3_COPY_1_5000
Perfect score: 5000
Sequence: 1 ctgggattggacacactt.....caagctgtcatcatatgac 5000

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: qb_estl:*
10: qb_est2:*
11: qb_hlc:*
12: qb_est3:*
13: qb_est4:*
14: qb_est5:*
15: em_estfun:*
16: em_estom:*
17: qb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	229	4.6	666	10	AV654588 AV654588
2	215	4.3	336	14	R07567 yeg7h07.r1
3	203.4	4.1	461	14	R07514 yeg7h07.s1
4	199.4	4.0	353	10	AV652896 AV652896
5	191	3.8	304	14	H63293 yr49e03.r1
6	190.6	3.8	444	14	R91117 yp93g09.r1

7	189.4	3.8	456	12	BF654744
8	188.6	3.8	454	14	H49223
9	183.4	3.7	529	13	BI343595
10	175.4	3.5	608	10	AV655576
11	173.8	3.5	529	9	AI303081
12	170	3.4	393	9	AA504868
13	167	3.2	597	10	AV650465
14	159	3.2	687	17	AQ878410
15	152.6	3.1	431	9	AI328320
16	152.6	3.1	475	14	W99028
17	152.2	3.0	320	9	AA343990
18	150.8	3.0	701	13	BI219735
19	145.2	2.9	419	9	AI025149
20	145.2	2.9	422	9	AI023507
21	140.6	2.8	419	14	T89505
22	134.2	2.7	420	14	H94827
23	129.8	2.6	558	17	AQ518926
24	128.4	2.6	277	14	T81776
25	128.4	2.6	550	17	AQ533551
26	125.2	2.5	324	10	AW085134
27	124.4	2.5	288	14	T77573
28	123.2	2.5	788	10	AV695776
29	123	2.5	499	14	H54245
30	121.8	2.4	592	17	AO582455
31	120	2.4	750	10	AV695796
32	114.8	2.3	440	17	AQ074543
33	112	2.2	620	9	AA237452
34	110	2.2	675	17	AG037548
35	109.8	2.2	375	10	AV690608
36	109	2.2	361	10	AV651158
37	108.8	2.2	275	9	AA572527
38	107.6	2.2	465	14	T89681
39	106.6	2.1	469	17	AQ009509
40	104	2.1	292	14	T82089
41	103.6	2.1	493	17	AQ292872
42	103.4	2.1	483	14	T77537
43	102.8	2.1	518	9	AI545087
44	102	2.0	423	9	AI051709
45	97.8	2.0	448	17	AQ220996

ALIGNMENTS

RESULT 1
AV654588
LOCUS AV654588 666 bp mRNA linear EST 15-JAN-2002
DEFINITION AV654588 GIC Homo sapiens CDNA clone GICDX08 3', mRNA sequence.
ACCESSION AV654588
VERSION AV654588.1 GI:9875602
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Euteleostomi; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 666)
Xiao,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xie,H., Liu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z., and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
CONTACT: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzyg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

```
FEATURES
  source
    Location/Qualifiers
      1..666
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="GLC/DX08"
        /clone_1lb="GLC"
        /tissue_type="corresponding non cancerous liver tissue"
        /dev_stage="adult"
        /lab_host="SOLR"
        /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      187 a      201 c      153 g      125 t
ORIGIN
Query Match      4.6%; Score 229; DB 10; Length 666;
Best Local Similarity 79.5%; Pred. No. 2e-59;
Matches 271; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 92 GAGCAGACCTGAGCAAGCCATGTGGTCCAGATGCTACCATGATGGAGCAGAGTT 151
DB 325 CAGCACACCTGAGCTAGCCCTGTGGTCCAGACTGCTACCATGATGGAGCAGAGCT 384
QY 152 ATCGAGGACGCTACTCCACACCTGCTCAGAGAAAGACCTGCCAAGCTGTCTATGA 211
DB 385 ACCGAGGACATACTTCCACACACACAGAGAAAGAGTGTGCTGTGCTATGTA 444
QY 212 CAGCAATCAACATATATGAGCACAGAAACTACCAATGCTGGCTGATCATGAACT 271
DB 445 CAGCACACCGGACAGACAGACCCAGAAACTACCAATGCTGGCTGATCATGAACT 504
QY 272 ACTGAGAAATCCAGATCTGTGGACGCTCTTATTTGATACAGAGGATCCGGGTGTA 331
DB 505 ACTGAGAGAAATCCAGATCTGTGGACGCTCTTATTTGATACAGAGGATCCGGGTGTA 564
QY 332 GGTGGAGTACGCAACCTGACGAGCTCAGACGAGGAGGAGTGGCGTGGCCCTC 391
DB 565 GGTGGAGTACGCAACCTGACGAGCTCAGACGAGGAGGAGTGGCGTGGCCCTC 624
QY 392 CGACTGTATACCCGGTTCAGACCTAGAGGCTCCTTCGAA 432
DB 625 CCCTGTGTGCTCCTGCTCCAGATGTAGAGACTCCTTCGAA 665
RESULT 2
LOCUS      R07567      336 bp      mRNA      linear      EST 05-APR-1995
DEFINITION yae97h07.1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:125725 5' similar to gb:X06290 APOLIPOPROTEIN(A) PRECURSOR
(HUMAN);, mRNA sequence.
ACCESSION  R07567
VERSION    R07567.1 GI:759490
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 336)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maira,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            R., Williamson,A., Wohlmann,P. and Wilson,R.
            The WashU-Merck EST Project
            Unpublished (1995)
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@wustl.wustl.edu
            Insert Size: 872
            High quality sequence stops: 327 Source: IMAGE Consortium, LNL. This
            clone is available royalty-free through LNL; contact the IMAGE
            Consortium (info@image.lnl.gov) for further information.
```

```
Insert Length: 872      Std Error: 0.00
Seq primer: M3RPI
High quality sequence stop: 327.
FEATURES
  source
    Location/Qualifiers
      1..336
        /organism="Homo sapiens"
        /db_xref="GDB:478270"
        /db_xref="taxon:9606"
        /clone="IMAGE:125725"
        /clone_1lb="Soares fetal liver spleen INFLS"
        /sex="male"
        /dev_stage="20 week post conception fetus"
        /lab_host="DH10B (ampicillin resistant)"
        /note="Organ: Liver and Spleen; Vector: p7773D (Pharmacia)
            with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
            1st strand cDNA was primed with a Pac I - Oligo(dT) primer
            15' AACGGAAGATTAATTAAGATCTTTTATTTTATTTTATTTTATTTTATTTT
            double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Pac I and cloned into the Pac I
            and Eco RI sites of the modified p7773 vector. Library
            went through one round of normalization. Library
            constructed by Bento Soares and M.Fallina Bonaldo."
BASE COUNT      93 a      92 c      75 g      68 t
ORIGIN
Query Match      4.3%; Score 215; DB 14; Length 336;
Best Local Similarity 82.1%; Pred. No. 2.8e-55;
Matches 256; Conservative 0; Mismatches 52; Indels 4; Gaps 1;
QY 193 CAAGCTTGATCTATGACACCATCAATCAATATAGACCAAGAAATACCAAT 252
DB 1 CAAGCTTGATCTATGACACCATCAATCAATATAGACCAAGAAATACCAAT 60
QY 253 GCTGGCTGATCATGTAACACTGACAGATCCAGATGCTGTGGACGCTCCTATTGTAT 312
DB 61 GCTGGCTGATCATGTAACACTGACAGATCCAGATCTCTGTGNAAGCCCTTGTGAC 120
QY 313 ACAGAGGATCCGGTGTACAGTGGAGTACTGCAACTGACCAATGCTCAGACGAG 372
DB 121 ACAGATGATCCCAAGTCAAGTGGAGTACTGCAACTGACCAATGCTCAGACGAG 180
QY 373 GGGACGCGCTGCGCGCTGCACTGTATACCGCGGTTCACAGCTAGAGGCTCCTCG 432
DB 181 GGGACGCGCTGCGCGCTGCACTGTATACCGCGGTTCACAGCTAGAGGCTCCTCG 240
QY 433 CAAGCACCAGCTGAGCAAGAGCCCTGGGGT---GCAGAGTGTCAACATGTAATGACA 488
DB 241 CCAGCACCAGCTGAGCAAGAGCCCTGGGGT---GCAGAGTGTCAACATGTAATGACA 300
QY 489 GAGTTATCGAGG 500
DB 301 GAGTTATCGAGG 312
RESULT 3
LOCUS      R07514/C      461 bp      mRNA      linear      EST 05-APR-1995
DEFINITION yae97h07.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:125725 3' similar to gb:X06290 APOLIPOPROTEIN(A) PRECURSOR
(HUMAN);, mRNA sequence.
ACCESSION  R07514
VERSION    R07514.1 GI:759437
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 461)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maira,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            R., Williamson,A., Wohlmann,P. and Wilson,R.
            The WashU-Merck EST Project
```


HE3293	HE3293	304 bp	MRNA	linear	EST 11-OCT-1996
LOCUS	YR49603.r1	Soares fetal liver spleen	INFLU	HOMO sapiens	CDNA clone
DEFINITION	IMAGE:208656.5, similar to gb.X05199	PLASMINOGEN	PRECURSOR	(HUMAN	
ACCESSION	HE3293				
VERSION	HE3293.1	GI:1018094			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 304)				
AUTHORS	Hillier, L., Clark, N., Dubouque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kuwab, T., Le, M., Lennon, G., Harris, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Treveskts, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.				
TITLE	The Mashu-Merck EST Project				
JOURNAL	Unpublished (1995)				
COMMENT	Contact: Wilson RK				

```

Insert size: 2912
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert length: 2912      Std Error: 0.00
Seq primer: M13mp1
High quality sequence stop: 1.
Location/Qualifiers

```

```

2:3001
/organism="Homo sapiens"
/db_xref="GDB:3777767"
/db_xref="taxon:9606"
/clone IMAGE:208636"
/clone.lib="Soares fetal liver spleen INFLS"
/dev="male"
/dev.stage="20 week-post conception fetus"
/lab.host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: p7773D (Pharmacia)
with a modified polylinker; Site.1: Pac I; Site.2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGCAAGAAATTAATTAAGAATCTTTTCTTTTCTTTTCTTTTCTTTT
3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p7773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patricia Bonaldo."

```

	Query Match	3.8%	Score 191	DB 14	Length 304	
	Best Local Similarity	75.4%	Pred. No. 7.1e-48			
	Matches 227	Conservative 0	Mismatches 74	Indels 0	Gaps 0	
Qy	108	AAGCATGTGGTCCAGGATTCGTACCATGATGATGACAGATATATGAGGACACGTAATC	167			
Db	4	AAGCCCTGTGTCCAGAGACTGCTACCATGTGTATGACAGAGCTACCGAGGACATCTCTC	63			
Qy	168	CACCACTGTCAACGAGGAACTGCCAAGCTTGGTCATCTATGACACACATCAACATTA	227			
Db	64	CANACACCACACAGGTAGAGATNTCAGTCTTGGTCATCTATGACACACACACGACCA	123			
Qy	228	TAGAGACAGAGAAATATACCAAAATGCGTGTATCTATGAACATACGACGAAATACAG	287			
Db	124	GANACCCACGAAATATACCGAATATGTTGGCTGTACANTGATACCTACGAGGNNTCAG	183			

Qy	288	TGCGTGTGGCAGCTCTTATTTGTTATATACAGAGGATCCCGGTGTACAGGGGAGGATCTGCAC	347
Qy	184	TGCGGATTAAGGCCCTCGTNTTTTACACAGACCCCGAGCGTACAGGTGGGATCTGCAC	243
Db	348	CGTGCAGCATGTGTCAGACGAGAGGGACATGCGCGTGGCGCTCCGACCTTATACCCCGGT	407
Qy	244	CGTGAATAAATGTCTCAGAGACAGAAAGCAGTGTGTAGACACTTCGCGCTGTTCCTGCT	303
Db	408	T	408
Qy	304	T	304

RESULT	6
R91117	
LOCUS	
DEFINITION	444 bp mRNA linear EST 25-AUG-1995 YP9309.r1 Soares fetal liver spleen INFES Homo sapiens cDNA clone IMAGE:135040.5 similar to gb:X05199 PLASMINOGEN PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION	R91117
VERSION	R91117.1 GI:958657
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 444)
AUTHORS	Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman .M., Hultman,M., Kucaba,T., Le,M., Lemonon,G., Marie,M., Parsons,J., Rifkin,L., Rolfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston .R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE	The WashU-Merck EST Project
JOURNAL	Unpublished (1995)
COMMENT	Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estelw@wustl.edu
Insert Size: 1916
High quality sequence stops: 310
Source: IMAGE Consortium, LNU.
This clone is available royalty-free through LNU ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1916 Std Error: 0.00
Seq Primer: M13R1
High quality sequence stop: 310.
Location/Qualifiers
1..444

```

/db.xref="taxon:9606"
/clone="IMAGE:195040"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistently)"
/note="Organ: Liver and Spleen; Vector: pT713D (Pharmacia) with a modified polylinker; Site.1: Pac I; Site.2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer 15', AACTGGAGAGATTAAATTAAGATCCCTTTTTTTTTTTTTTTT 3'), double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Patima Bonaldo."
BASE COUNT      127 a      130 c      104 g      81 t
ORIGIN
Query Match      3.8%; Score 190.6; DB 14; Length 444;
Best Local Similarity 81.5%; Pred.No. 1.2e-47;
Matches 220; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
```

Query Match	3.88;	Score 189.4;	DB 12;	Length 456;
Best Local Similarity	81.28;	Pred. No. 2.8e-47;		

```

/organism="Homo sapiens"
/db_xref="CDB:3798132"
/db_xref="taxon:9606"
/clone="IMAGE:274351"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pRTT3D (Pharmacia) with a modified polylinker; Site.1: Pac I - Site.2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACGCGAGCAATTAACTTAAGACATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I

```



```
FEATURES      High quality sequence stop: 385.
SOURCE        Location/Qualifiers
              1..393
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_image:839703"
              /clone_lib="Stratagene fetal retina 937202"
              /sex="mixed"
              /lab_host="SOLR (kanamycin resistant)"
              /note="Vector: pBluescript SK-, Site_1: EcoRI, Site_2:
              XhoI; Cloned unidirectionally. Primer: Oligo dT, pooled
              retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
              Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3',
              adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'"

BASE COUNT    102 a 112 c 90 g 89 t
ORIGIN
Query Match    3.4%; Score 170; DB 9; Length 393;
Best Local Similarity 99.5%; Pred. No. 2.6e-41;
Matches 181; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 255 TGGCTGATCATGAACTACTGAGAAATCCAGATGCTGTGCGAGCTCCTTATGTATAC 314
      |||||||
Db 226 TGGCTGATCATGAACTACTGAGAAATCCAGATGCTGTGCGAGCTCCTTATGTATAC 167

QY 315 GAGGATCCCGGTGTGAGGTGGAGTACGCAACCTGACGCAATGCTCAGAGAGAG 374
      |||||||
Db 166 GAGGATCCCGGTGTGAGGTGGAGTACGCAACCTGACGCAATGCTCAGAGAGAG 107

QY 375 GACTCGCGTCCGCTCCGACTGTACCCCGGTTCCAGAGCTAGAGGCTCTTCGAGACA 434
      |||||||
Db 106 GACTCGCGTCCGCTCCGACTGTACCCCGGTTCCAGAGCTAGAGGCTCTTCGAGACA 48

QY 435 AG 436
      ||
Db 47 AG 46

RESULT 13
LOCUS      AV650465 687 bp mRNA linear EST 15-JAN-2002
DEFINITION AV650465 GIC Homo sapiens cDNA clone GICCFB02 3', mRNA sequence.
ACCESSION  AV650465
VERSION     AV650465.1 GI:9871479
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 687)
AUTHORS     Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
             Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
             Shen,K., Lu,G., Fu,G., Zhong,W., Xu,S., Gu,W., Huang,W., Zhao,X.,
             Hu,G., Gu,J., Chen,Z., and Han,Z.
             Insight into hepatocellular carcinogenesis at transcriptome level
             by comparing gene expression profiles of hepatocellular carcinoma
             with those of corresponding noncancerous liver
             Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
COMMENT     Contact: Zegang Han
             Chinese National Human Genome Center at Shanghai
             351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
             201203, P. R. China
             Tel: 86-21-50801919(ex.45)
             Fax: 86-21-50801922
             Email: hanzegang.sh.cn
             This clone is available at CHGC in Shanghai.

FEATURES
SOURCE      1..687
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_image="GICCFB02"
             /clone_lib="GIC"

FEATURES      /tissue_type="corresponding non cancerous liver tissue"
SOURCE        /dev_stage="Adult"
              /lab_host="SOLR"
              /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
              XhoI"

BASE COUNT    185 a 174 c 181 g 138 t 9 others
ORIGIN
Query Match    3.3%; Score 167; DB 10; Length 687;
Best Local Similarity 77.8%; Pred. No. 3e-40;
Matches 214; Conservative 0; Mismatches 60; Indels 1; Gaps 1;

QY 163 TACTCCACCTAGTTCACAGAGAGACCTGCGCAAGCTGTGTCATGTATGACACCAATCAA 222
      |||||||
Db 8 TCTCTCACACACACACACAGAGAGAGTGTCACTCTTGTGTCATGTATGACACCAACCGG 67

QY 223 CATATATGAGACCAAGAAACTACCCAAATGCTGCTTATGATCATGACTGAGAAAT 282
      |||||||
Db 68 CACCAAGAGAGCCCAAGAAACTACCCAAATGCTGCTTATGATCATGACTGAGAAAT 127

QY 283 CCAGATGCTGTGGAGCTCCTTATGTATGAGAGGATCCGGTGTGAGTGGAGTAC 342
      |||||||
Db 128 CCAGATGCTGTGGAGCTCCTTATGTATGAGAGGATCCGGTGTGAGTGGAGTAC 187

QY 343 TGCAAACCTG-ACGCAATGCTCAGACGAGAGGAGTGGCTGGCGCTCCGACTGTAC 401
      |||||||
Db 188 TGCAACTGAAAAAATGCTCAGAGACAGAGAGGATGTGTGTAGCACTCCGACTGTGT 247

QY 402 CCCGTTCCAGGCTAGAGGCTCCTTCGAGCAAG 436
      |||||||
Db 248 CCGTCTTCAGATGTAGAGACTCTTCGAGCAAG 282

RESULT 14
LOCUS      A0878410 597 bp DNA linear GSS 09-NOV-1999
DEFINITION HS_3069_A1.B10.T7C CTR Approved Human Genomic Sperm Library D Homo
             sapiens genomic clone Plate-3069 Col-19 Row-C, DNA sequence.
ACCESSION  A0878410
VERSION     A0878410.1 GI:6309877
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 597)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
             Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
             Hood,L.
             Sequence-tagged connectors: A sequence approach to mapping and
             scanning the human genome
             Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
             High Throughput Sequencing Center
             University of Washington
             401 Queen Anne Avenue North, Seattle, WA 98109, USA
             Tel: (206) 616-3618
             Fax: (206) 616-3887
             Email: jwallace@u.washington.edu
             Clones may be purchased from Research Genetics (info@resgen.com).
             BAC end Web Server: http://www.hisc.washington.edu
             Plate: 3069 row: C column: 19
             Seq primer: T7
             Class: BAC ends
             High quality sequence stop: 597.

FEATURES
SOURCE      1..597
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_image="Plate=3069 Col=19 Row=C"
             /clone_lib="CTR Approved Human Genomic Sperm Library D"
             /sex="male"
```

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 154 a 134 c 132 g 166 t 11 others

ORIGIN

Query Match 3.2%; Score 159; DB 17; Length 597;
Best Local Similarity 98.1%; Pred. No. 8.2e-38;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 435 AGCACCGACTGAGCAAGCGCTGGGGGTGTCAGAGTGTACCATGATGATGAGAGAGTTA 494
|||||
Db 505 AGCAGNAGCTAGCAAGAGCGTGGGTGTCAGAGTGTACCATGATGATGAGAGAGTTA 446
|||||
QY 495 TCGAGGACATATCTCCACCACTGTACAGAGAAAGCTGCCAAGCTTGTATCTATGAC 554
|||||
Db 445 TCGAGGACATATCTCCACCACTGTACAGAGAAAGCTGCCAAGCTTGTATCTATGAC 386
|||||
QY 555 ACCAGACTGCGATAGTGGAGCCCGAGATATCTACCCAAATGC 596
|||||
Db 385 ACCAGACTGCGATAGTGGAGCCCGAGATATCTACCCAAATGC 344
|||||

RESULT 15 431 bp mRNA linear EST 18-MAR-1999
A1529320 u162308 y1 Sugano mouse liver mlia Mus musculus cDNA clone
LOCUS IMAGE:1886966 5' similar to gb:U05199 PLASMINOGEN PRECURSOR (HUMAN
DEFINITION); gb:U04766 Mouse plasminogen mRNA, complete cds (MOUSE);, mRNA
sequence.
ACCESSION A1529320
VERSION A1529320.1 GI:4443455
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 431)
Marras, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
'B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Rilter
'E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The Mashu-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Maria M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:971290
Seq primer: custom primer used.
Location/Qualifiers
1. .431
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1886966"
/clone_1db="Sugano mouse liver mlia"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: liver; Vector: pME18S-FL3; Site:1: DraIII
(CACTGCTG); Site:2: DraIII (CACCAGTGG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGCTG, 3' site CACCAGTGG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library

constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTTAAAGCTGCG and 3' end
primer CGACCTGACAGCTCGAGCACA."

BASE COUNT 108 a 125 c 111 g 87 t

ORIGIN

Query Match 3.1%; Score 152.6; DB 9; Length 431;
Best Local Similarity 72.9%; Pred. No. 6.5e-36;
Matches 196; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 95 CAGCACCTGAGCAAGAGCCATGTGTCCAGATGTGCTACCATGATGATGAGAGAGTATAC 154
|||||
Db 60 CACCGAGAGAGCAAAACACTGTGTGTCAGGATCTTCCAGAGGAGGAGGAGAGTATAC 119
|||||
QY 155 GAGGACGCTACTCCACCACTGTACAGAGAAAGCTGCCAAGCTTGTATCTATGACAC 214
|||||
Db 120 GGGGTACATCGTCCACTACCATCAAGGAGAGAGTGGAGTCTGAGATGTTTC 179
|||||
QY 215 CACATCAACATATATAGGACACAGAGAAAGTACCAATGCTGGCTTGTATGATGAACTACT 274
|||||
Db 180 CACACAGGCATTCGAAAGACCCAGAGAACTTCCAGATGCTGGCTTGGAGATGAACTACT 239
|||||
QY 275 GCAGGATCCAGATGCTGTGGAGCTCCTTATTTATACAGAGGATCCGGTGTACAGT 334
|||||
Db 240 GCAGGATCCAGATGCTGTGGAGCTCCTTATTTATACAGAGGATCCGGTGTACAGT 299
|||||
QY 335 GGGAGTACTGCAACCTGACGCAATGCTCA 363
|||||
Db 300 GGGAACTACTGCAACCTGAAAGGGGTGCTCA 328
|||||

Search completed: March 6, 2003, 15:19:13
Job time : 6007.5 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 01:17:11 ; Search time 12208 Seconds

(without alignments)
11919.563 Million cell updates/sec

Title: US-09-923-515-3_COPY_8939_13938

Perfect score: 5000
Sequence: 1 ctccgaccttaccgccglt.....ctgcactatttgatttga 5000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pla:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5000	100.0	13938	9 HSAALIPOA	X06290 Human mRNA
2	3747.2	74.9	4510	9 MACAPOA	J04635 Rhesus monk
3	1237	24.7	2753	6 I45623	I45623 Sequence 1
4	1235.4	24.7	2732	9 HSEPMGR	X05199 Human mRNA
5	1235.4	24.7	2753	6 A22096	A22096 plasmidogen
6	1123.8	24.7	1784	9 H0MPWGM	K02922 Human plasm
7	1128.8	24.6	2679	6 I06212	I06212 Sequence 14
8	1178.8	23.6	2711	9 MACPEPSC	J04697 Rhesus monk
9	1110.6	22.2	2296	6 AR105749	AR105749 Sequence
10	1110.6	22.2	2296	6 I14104	I14104 Sequence 18
11	1106.4	22.1	2497	6 AX448883	AX448883 Sequence
12	1104.8	22.1	2497	6 AR037325	AR037325 Sequence
13	1104.8	22.1	2497	6 AR082437	AR082437 Sequence
14	1104.8	22.1	2497	6 AR085163	AR085163 Sequence
15	1104.8	22.1	2497	9 HUMPLASM	M74220 Human plasm
16	1089.8	21.8	2433	6 AX463622	AX463622 Sequence
17	989.4	19.8	1649	9 AF029691	AF029691 Papio ham
18	880.2	17.6	1272	9 AF029692	AF029692 Papio ham
19	868.6	17.4	2770	4 BPTLASM1N	X79402 B. taurus mk
20	858.6	17.2	2041	9 HSU19517	U19517 Human (apoa
21	857.6	17.2	2771	10 BC014773	BC014773 Mus muscu
22	851.8	17.0	2720	10 M0SPPLGN	J04766 Mouse plasm
23	845.4	16.9	1047	6 AX463626	AX463626 Sequence
24	830.8	16.6	2737	10 RNO242649	AJ242649 Rattus no
25	775.4	15.5	2745	4 EFU33171	U33171 Erinaceus e
26	740.8	14.8	1712	9 HSU19518	U19518 Human (apoa
27	693	13.9	2832	4 AF012297	AF012297 Macropus
28	613	12.3	750	6 AX463624	AX463624 Sequence
29	557.8	11.2	690	6 A37078	A37078 Sequence 1
30	557.8	11.2	690	6 I51668	I51668 Sequence 1
31	410	8.2	135751	9 AL596089	AL596089 Human DNA
32	408.4	8.2	271762	2 AL645523	AL645523 Homo sapi
33	325.4	6.5	171363	2 AC084863	AC084863 Papio cyn
34	311.6	6.2	1353	6 AX117488	AX117488 Sequence
35	306.4	6.1	1236	6 AX329649	AX329649 Sequence
36	306.4	6.1	1236	6 AX336098	AX336098 Sequence
37	306.4	6.1	1236	6 AX336421	AX336421 Sequence
38	306.4	6.1	1236	6 HUMPLG24	M34276 Human plasm
39	306.4	6.1	2589	9 AK096980	AK096980 Homo sapi
40	306.4	6.1	162083	9 AL591069	AL591069 Human DNA
41	306.4	6.1	163599	9 HSDJ81D8	AL109933 Homo sapi
42	306.4	6.1	177878	2 AC010893	AC010893 Homo sapi
43	306.4	6.1	200402	2 AC040893	AC040893 Homo sapi
44	306.4	6.1	271762	2 AL645523	AL645523 Homo sapi
45	283	5.7	435	9 S79621	S79621 apolipoprot

ALIGNMENTS

RESULT 1
LOCUS HSAALIPOA 13938 bp mRNA linear PRI 30-MAR-1995
DEFINITION Human mRNA for apolipoprotein(a).
ACCESSION X06290.1 GI:28619
VERSION X06290.1 GI:28619
KEYWORDS apolipoprotein; apolipoprotein A; glycoprotein; lipoprotein;
plasmaogen; serine protease.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Xenotopia; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 13938)
AUTHORS McLean, J.W.
TITLE Direct Submission

QY 361 TTCCAGGCTTAGAGGCTCTTCCGAACACGACGACTGAGAGGCGCTGGGGTCCAG 420
|||||
Db 9299 TTCCAGGCTTAGAGGCTCTTCCGAACACGACGACTGAGAGGCGCTGGGGTCCAG 9358
QY 421 AGTGTACACGAGTAAATGAGACAGATTATGAGAGGACATATCCACACTGTCTACTGGAA 480
|||||
Db 9359 AGTGTACACGAGTAAATGAGACAGATTATGAGAGGACATATCCACACTGTCTACTGGAA 9418
QY 481 GAACCTGCAAGCTTGGTCAATATGACACACACTGCAATAGTGGAGCCCAAGATATCT 540
|||||
Db 9419 GAACCTGCAAGCTTGGTCAATATGAGACACACACTGCAATAGTGGAGCCCAAGATATCT 9478
QY 541 ACCCAATGCTGGCTTGATCATGAACTACTGACGAATCCAGATGCTGTGGACGCTCTT 600
|||||
Db 9479 ACCCAATGCTGGCTTGATCATGAACTACTGACGAATCCAGATGCTGTGGACGCTCTT 9538
QY 601 ATTGTTATACGAGGATCCGGGTGTCAAGTGGAGATGACAACTGAGCAATGCTCAG 660
|||||
Db 9539 ATTGTTATACGAGGATCCGGGTGTCAAGTGGAGATGACAACTGAGCAATGCTCAG 9598
QY 661 ACCGAGAGGAGCTGCGGTGCGGCTCCAGCTTACCCCGGTTCCAAAGCTTAGAGGCTC 720
|||||
Db 9539 ACCGAGAGGAGCTGCGGTGCGGCTCCAGCTTACCCCGGTTCCAAAGCTTAGAGGCTC 9658
QY 721 CTTCCGACACAGCACCCGACAGAGAGGCGCTGGGGTGCAGAGGCTGTACCCAGGTAATG 780
|||||
Db 9659 CTTCCGACACAGCACCCGACAGAGAGGCGCTGGGGTGCAGAGGCTGTACCCAGGTAATG 9718
QY 781 GACAGAGTTATCGAGGACATATCTCCACACTGTCACTGGAAAGAACCTCCAAAGCTTG 840
|||||
Db 9719 GACAGAGTTATCGAGGACATATCTCCACACTGTCACTGGAAAGAACCTCCAAAGCTTG 9778
QY 841 CATCTATGACACCCACACTCGCATATGTTGGAGCCCAAGATCTCCCAATGCTGGTTGA 900
|||||
Db 9779 CATCTATGACACCCACACTCGCATATGTTGGAGCCCAAGATCTCCCAATGCTGGTTGA 9838
QY 901 TCATGAACTACTGACGAATCCAGATGCTGTGGCAAGCTCTTATTTATATACGAGGATC 960
|||||
Db 9839 TCATGAACTACTGACGAATCCAGATGCTGTGGCAAGCTCTTATTTATATACGAGGATC 9898
QY 961 CCGGTGTAGGAGGAGAGTACGAACTGACGGCAATGCTCAGACGAGAGAGGAGCGCCG 1020
|||||
Db 9899 CCGGTGTAGGAGGAGAGTACGAACTGACGGCAATGCTCAGACGAGAGAGGAGCGCCG 9958
QY 1021 TCGCGCTTCGCACTGTTAACCCGGGTTCCAAAGCTTAGAGGCTCTTCCGAACAAGCACGA 1080
|||||
Db 9959 TCGCGCTTCGCACTGTTAACCCGGGTTCCAAAGCTTAGAGGCTCTTCCGAACAAGCACGA 10018
QY 1081 CTGAGCAGAGGCGTGGGGTGGAGAGTGTACACGCGTATATGACAGAGTTATCGAGGCA 1140
|||||
Db 10019 CTGAGCAGAGGCGTGGGGTGGAGAGTGTACACGCGTATATGACAGAGTTATCGAGGCA 10078
QY 1141 CATACTCCACACACTGTCACTGGAAGAACCTGCCAAGCTTGTCATCTATGACACCACT 1200
|||||
Db 10079 CATACTCCACACACTGTCACTGGAAGAACCTGCCAAGCTTGTCATCTATGACACCACT 10138
QY 1201 GCGAATAGTCGACCCCAATAATCTACCAATGCTGGCTGATCAATGACTACTGTGAGGA 1260
|||||
Db 10139 GCGAATAGTCGACCCCAATAATCTACCAATGCTGGCTGATCAATGACTACTGTGAGGA 10198
QY 1261 ATCCAGATCTGTGGCAGCCCTTATTGTTATATACGAGGATCCAGATGTCAGGTGGAGT 1320
|||||
Db 10199 ATCCAGATCTGTGGCAGCCCTTATTGTTATATACGAGGATCCAGATGTCAGGTGGAGT 10258
QY 1321 ACTGGAACCTGACAAATGCTCAGACGAGAGGAGTGGCGGTCCCAACTATATTA 1380
|||||
Db 10259 ACTGGAACCTGACAAATGCTCAGACGAGAGGAGTGGCGGTCCCAACTATATTA 10318
QY 1381 CCCCGATTCCAAAGCTTAGAGGCTCTTCTGAAACAAGCACCACTGAGCAAAAGGCTGGGG 1440
|||||
Db 10319 CCCCGATTCCAAAGCTTAGAGGCTCTTCTGAAACAAGCACCACTGAGCAAAAGGCTGGGG 10378

QY 1441 TGCAGAGTGTCTACACCGAAATGACAGAGTATATCAAGGCACATCTTCACTACTGTCA 1500
|||||
Db 10379 TGCAGAGTGTCTACACCGAAATGACAGAGTATATCAAGGCACATCTTCACTACTGTCA 10438
QY 1501 CAGGAAGAACCTGCAAGCTTGCTATCTATGACACCACTGCGATATGTCGAGCCCGAG 1560
|||||
Db 10439 CAGGAAGAACCTGCAAGCTTGCTATCTATGACACCACTGCGATATGTCGAGCCCGAG 10498
QY 1561 CATACTACCCAAATGCTGGCTTGATCAAGAACTACTGCGCAAAATCCAGATCTGTGGCAG 1620
|||||
Db 10499 CATACTACCCAAATGCTGGCTTGATCAAGAACTACTGCGCAAAATCCAGATCTGTGGCAG 10558
QY 1621 CCCCTGGTGTATATCAACAGATCCCACTGTGACAGTGGAGGATCTGCACTTACACGAT 1680
|||||
Db 10559 CCCCTGGTGTATATCAACAGATCCCACTGTGACAGTGGAGGATCTGCACTTACACGAT 10618
QY 1681 GCTCAGATCAGAAATGAGACTGCTTCGCTCCGCAATGTTATCTGGCTCCAAAGCTTGA 1740
|||||
Db 10619 GCTCAGATCAGAAATGAGACTGCTTCGCTCCGCAATGTTATCTGGCTCCAAAGCTTGA 10678
QY 1741 AGGCTTTTATTAACAACAGCACTGACTGAGAAACCCCGGGGTACAGACTGCTTACTACC 1800
|||||
Db 10679 AGGCTTTTATTAACAACAGCACTGACTGAGAAACCCCGGGGTACAGACTGCTTACTACC 10738
QY 1801 ATTATGACAGAGTTATCCAGAGCACATCTCCACACTGTACACGAGAACTTGGCAAG 1860
|||||
Db 10739 ATTATGACAGAGTTATCCAGAGCACATCTCCACACTGTACACGAGAACTTGGCAAG 10798
QY 1861 CTTGTGATCTATGACACCACACAGCATATGTGGAGCCCGAATACTCCAAATGCTG 1920
|||||
Db 10799 CTTGTGATCTATGACACCACACAGCATATGTGGAGCCCGAATACTCCAAATGCTG 10858
QY 1921 GCGTACACAGAACTACTGACAGAAATCCAGATGCTGAGATTCGCGCTTGGTTATACCA 1980
|||||
Db 10859 GCGTACACAGAACTACTGACAGAAATCCAGATGCTGAGATTCGCGCTTGGTTATACCA 10918
QY 1981 TGGATCCCAAGTGTCAAGTGGAGTACTGCAACCTGACCAATGCTTGGTACAGAAATCAA 2040
|||||
Db 10919 TGGATCCCAAGTGTCAAGTGGAGTACTGCAACCTGACCAATGCTTGGTACAGAAATCAA 10978
QY 2041 GTGTCTTGTCACTGTCAAGGTTGTCGCCAGATCCAAAGCACAGAGGCTTCTTGTGAAGAAG 2100
|||||
Db 10979 GTGTCTTGTCACTGTCAAGGTTGTCGCCAGATCCAAAGCACAGAGGCTTCTTGTGAAGAAG 11038
QY 2101 CACCAACGAGCAAGCCCCCGGGTCCAGAGATGCTTACATGCTGATGAGACAGATTAATC 2160
|||||
Db 11039 CACCAACGAGCAAGCCCCCGGGTCCAGAGATGCTTACATGCTGATGAGACAGATTAATC 11098
QY 2161 GAGGCTCATCTCTATACACTGTACAGAGAGGACATGCAAGTCTTGGTCTTATGACAC 2220
|||||
Db 11099 GAGGCTCATCTCTATACACTGTACAGAGAGGACATGCAAGTCTTGGTCTTATGACAC 11158
QY 2221 CACACTGCACTCAGAGGACACAGAAATTTATCCAAATGCTGGCGCTTACACAGAACTACT 2280
|||||
Db 11159 CACACTGCACTCAGAGGACACAGAAATTTATCCAAATGCTGGCGCTTACACAGAACTACT 11218
QY 2281 GCAAGAAATCCAGATGCTGATTAAGTCTTGGTATATACCATGATCCCAATGTAGAT 2340
|||||
Db 11219 GCAAGAAATCCAGATGCTGATTAAGTCTTGGTATATACCATGATCCCAATGTAGAT 11278
QY 2341 GGGAGTACTGCAACTGACACAAATGTCAGATGACAAATCAAGTGTCTTGGCAGCTCCA 2400
|||||
Db 11279 GGGAGTACTGCAACTGACACAAATGTCAGATGACAAATCAAGTGTCTTGGCAGCTCCA 11338
QY 2401 CCGGTGTTCTTGAACAAGACCAAGGAGCAAAAGCCCAAGTCCAGAGCTGTACTACAT 2460
|||||
Db 11339 CCGGTGTTCTTGAACAAGACCAAGGAGCAAAAGCCCAAGTCCAGAGCTGTACTACAT 11398
QY 2461 GTGATGACAGAGTTATGAGAGCTATCTCCACACTGTTTACAGAGAGACATGTCAAT 2520
|||||
Db 11399 GTGATGACAGAGTTATGAGAGCTATCTCCACACTGTTTACAGAGAGACATGTCAAT 11458
QY 2521 CTTGGTCTCTATGACACCACTGGCATCAGAGAAACCAAGAAATACTACCCAAATGCTG 2580

Db	11459	CTTGCTCTCTATGACACCACTGGCATCAGAGAACACAGAAATACCAAAAGG	1151E
OY	2581	GCCTGACCGAGAACTACTGCAAGAAATCCAGATGCTGAGATTCGGCCTTGTTATACCA	2640
Db	11519	GCCGACACAGAACTACTGACAGAAATCCAGATGCTGAGATTCGGCCTTGTTATACCA	11578F
OY	2641	TGATATCCAGTGCAGATGGGAGTACTGCAACTCTGACCAATGTCACGATAGTAATCAA	2700
Db	11579	TGGATCCCATGTGCAGATGGGAGTACTGCAACTCTGACCAATGTCACGATAGTAATCAA	11638B
OY	2701	CTCTCTCACAACTCCACGGTGTGCCAGTTCCAAGCACAGAGCTTCCTTCTGAAGAAG	2760
Db	11639	CTCTCTCACAACTCCACGGTGTGCCAGTTCCAAGCACAGAGCTTCCTTCTGAAGAAG	11698B
OY	2761	CACCAATGAAAAACACACTGGGGATCCAGAACTGCTACCGAGTATGAGACAGATTATC	2820
Db	11699	CACCAACTGAAAAACACACTGGGGATCCAGAACTGCTACCGAGTATGAGACAGATTATC	11758F
OY	2821	GAGGCACATCTCCACACTATCAGAGAAACATGTCACTCTTGGTGTCTATGACAC	2880
Db	11759	GAGGCACATCTCCACACTATCAGAGAAACATGTCACTCTTGGTGTCTATGACAC	11818B
OY	2881	CACATTCGCATGGAGATGCCATTATCTATCCAAATGCGGGCTACACAGAACTACT	2940
Db	11819	CACATTCGCATGGAGATGCCATTATCTATCCAAATGCGGGCTACACAGAACTACT	11878F
OY	2941	GCAGGAATCCAGATGCTGAGATTCGCCCTTGTTACACCATGATCCAGTCAAGT	3000
Db	11879	GCAGGAATCCAGATGCTGAGATTCGCCCTTGTTACACCATGATCCAGTCAAGT	11938B
OY	3001	GGGAGTACTCAACTGACACATGCTCCAGTGCAGATGCAGATGCTCTCACAACTCCA	3060
Db	11939	GGGAGTACTCAACTGACACATGCTCCAGTGCAGATGCAGATGCTCTCACAACTCCA	11998B
OY	3061	CAGTGGCCCGGTTCCAAACACANAGGCTCTTCTGAACAAGCACCTGAGAAAAACC	3120
Db	11999	CAGTGGCCCGGTTCCAAACACAGAGGCTCTTCTGAACAAGCACCTGAGAAAAAGCC	12058B
OY	3121	CTGTGTCTCAGGATTCCTTCACTACATGGTGTATGACGAGAGTTATCGAGCATATCTCCACCA	3180
Db	12059	CTGTGTCTCAGGATTCCTTCACTACATGGTGTATGACGAGAGTTATCGAGCATATCTCCACCA	12118B
OY	3181	CTGTGCACAGGAAGACCTCTCAATCTTGCTATCTATGATACACACTGSGCATCAGAGA	3240
Db	12119	CTGTGCACAGGAAGACCTCTCAATCTTGCTATCTATGATACACACTGSGCATCAGAGA	12178F
OY	3241	CCCCAGAAAATACCCMAATGCTGGCTCAGACCGAAGACTTACTGAGGAATCCAGATTCTG	3300
Db	12179	CCCCAGAAAATACCCMAATGCTGGCTCAGACCGAAGACTTACTGAGGAATCCAGATTCTG	12238B
OY	3301	GGAACACACCTGGTGTCAACAACCGAATCCGTGTGAGAGTGGGAGTACTGCAATCTGA	3360
Db	12239	GGAACACACCTGGTGTCAACAACCGAATCCGTGTGAGAGTGGGAGTACTGCAATCTGA	12298B
OY	3361	CACATCTCTCAGAAACAGAAATCAGGTGTCTAGAGACTCCCACTGTTGTTCAGTTCCAA	3420
Db	12299	CACATCTCTCAGAAACAGAAATCAGGTGTCTAGAGACTCCCACTGTTGTTCAGTTCCAA	12358B
OY	3421	GCATGAGGGCTCTTCTGAAAGACAGCAACTGAGCAAAACCCGTGTGTCGCCGCACTGGT	3480
Db	12359	GCATGAGGGCTCTTCTGAAAGACAGCAACTGAGCAAAACCCGTGTGTCCGCACTGGT	12418B
OY	3481	ACCATGGTAATGGCCAGAGATTATGAGGACCAATTCCTCCACCCTGTACAGAGAAAGCAT	3540
Db	12419	ACCATGGTAATGGCCAGAGATTATGAGGACCAATTCCTCCACCCTGTACAGAGAAAGCAT	12478F
OY	3541	GTCAAATCTTGATCATCAGACACCAACCGGCAATCAGAGACCCCAAGAAATACCCAA	3600
Db	12479	GTCAAATCTTGATCATCAGACACCAACCGGCAATCAGAGACCCCAAGAAATACCCAA	12538B
OY	3601	ATGATGGCTGCACATGAATCTACTGCAGGAATCCAGATGCCCATACAGGCCCTTGTTGTT	3660

Dd	12539	ATGATGCCCTGACATGACACTACTGTCAGGAATCCAGATGCCGATACAGGCCCTTGGTGT	12539
Qy	3661	TTACCATGAGACCCAGCATCAGGTGGAGTACTGCAACCTGACGCAGTGTCTGACAGACAG	3720
Dd	12599	TTACCATGAGACCCAGCATCAGGTGGAGTACTGCAACCTGACGCAGTGTCTGACAGACAG	12658
Qy	3721	AAGGACAGTGGTGGCCCTCCGACAGTGCATCAGAGTTCGAAAGCCTGAGGCGCTCTTTCG	3780
Dd	12659	AAGGACAGTGGTGGCCCTCCGACAGTGCATCAGAGTTCGAAAGCCTGAGGCGCTCTTTCG	12718
Qy	3781	AACAAGACTGATGTGTTGGGAATGGGAAGGATACCGGGGCAAGAGGCAACCATGTGA	3840
Dd	12719	AACAAGACTGATGTGTTGGGAATGGGAAGGATACCGGGGCAAGAGGCAACCATGTGA	12778
Qy	3841	CTGGGAGGCCATGCGCAGGAATGGGGCGCCAGAGGCCCATGAGACACAGACCTTCATTC	3900
Dd	12779	CTGGGAGGCCATGCGCAGGAATGGGGCGCCAGAGGCCCATGAGACACAGACCTTCATTC	12838
Qy	3901	CAGGAGCAAAATAATGGCGAGTCTGGAAAAAAATTACTGCCCTTACCCCTGATGTGACA	3960
Dd	12839	CAGGAGCAAAATAATGGCGAGTCTGGAAAAAAATTACTGCCCTTACCCCTGATGTGACA	12898
Qy	3961	TCATATGGCCCTGGTGTACACATGAAATCCAGAAACCTTTGACTGATGTGATCC	4020
Dd	12899	TCATATGGCCCTGGTGTACACATGAAATCCAGAAACCTTTGACTGATGTGATCC	12958
Qy	4021	CTCTCTGTGATCCTCTTCAATTTGATTGGGAAACCTCAAGTGAAGCCGACAGAAATATGC	4080
Dd	12959	CTCTCTGTGATCCTCTTCAATTTGATTGGGAAACCTCAAGTGAAGCCGACAGAAATATGC	13018
Qy	4081	CTGGAAGCATTTGAGGGGGGTGTGTGGCCACCAACATTCCTGGCCCTGGCAAGTACATC	4140
Dd	13019	CTGGAAGCATTTGAGGGGGGTGTGTGGCCACCAACATTCCTGGCCCTGGCAAGTACATC	13078
Qy	4141	TCAGACACAGGTTTGGAAAGCACTTCTGTGAGGACCTTAATATCCCAAGTGGGTGC	4200
Dd	13079	TCAGACACAGGTTTGGAAAGCACTTCTGTGAGGACCTTAATATCCCAAGTGGGTGC	13138
Qy	4201	TGACTGCTGTCACCTGCTTGAAGAAGTCCCTCAAGCCCTTCACTCTACAAAGTATCTGG	4260
Dd	13139	TGACTGCTGTCACCTGCTTGAAGAAGTCCCTCAAGCCCTTCACTCTACAAAGTATCTGG	13198
Qy	4261	GTGCACACCAAGAGTGAACCTTCGAATCTCATGTTCAGGAATATGAAGTGTCAAGCTGT	4320
Dd	13199	GTGCACACCAAGAGTGAACCTTCGAATCTCATGTTCAGGAATATGAAGTGTCAAGCTGT	13258
Qy	4321	TCTTGGAGCCCAACACAGAGATATTGCTCTGCTTAAAGCTAAGCAGGCGCCGCTATCA	4380
Dd	13259	TCTTGGAGCCCAACACAGAGATATTGCTCTGCTTAAAGCTAAGCAGGCGCCGCTATCA	13318
Qy	4381	CTGCAAAAGTAATGCGAGTGTCTGCGCATCCCAAGACTACAGTGTACCGCCAGACATC	4440
Dd	13319	CTGCAAAAGTAATGCGAGTGTGTCTGCGCATCCCAAGACTACAGTGTACCGCCAGACATC	13378
Qy	4441	AATGTTCATCACTGTGCTGGGGAAGAACCAAGTACCTTTTGGGACATGCGCTTCTCAAG	4500
Dd	13379	AATGTTCATCACTGTGCTGGGGAAGAACCAAGTACCTTTTGGGACATGCGCTTCTCAAG	13438
Qy	4501	AAGCCACGCTCTGTTATTATGGAATGAAGTGTGCAATCACTAATAGTAAATTGTGATC	4560
Dd	13439	AAGCCACGCTCTGTTATTATGGAATGAAGTGTGCAATCACTAATAGTAAATTGTGATC	13498
Qy	4561	AGCATTTGGCCAGAGGCACTGACAGTGTCCAGAGGTGACAGTGGAGGGGCTCTGGTTGCT	4620
Dd	13499	AGCATTTGGCCAGAGGCACTGACAGTGTCCAGAGGTGACAGTGGAGGGGCTCTGGTTGCT	13558
Qy	4621	TCGAGAGGCAAAATCAATTTTACAGAGTGCACCTGCTGGGGGCTTGGGTGACAGCC	4680
Dd	13559	TCGAGAGGCAAAATCAATTTTACAGAGTGCACCTTCTGGGGGCTTGGGTGACAGCC	13618
Qy	4681	CCAAATAGCCGTGTCTATGCTGTGTTCAAGGTTTGTACTGATTTGAGGGAATCA	4740
Dd	13619	CCAAATAGCCGTGTGTATGCTGTGTTCAAGGTTTGTACTGATTTGAGGGAATCA	13678

QY 1176 GCTTGGTCATCTATGACACACACTCCGATAGTGGACCCCGAGAACTACTACCCAATGCT 1235
|||||
Db 901 GCTTGGTCATCTATGAAACACACACTCTCTAGTGGACCCCGGAAACTACCCAATGCT 960
QY 1236 GGCTTGAATATAGTACTGAGAGAAATCCAGATCCTGTGGACGCCCTTATTTATTTAG 1295
|||||
Db 961 GGCTTGAATAGAACTACTGAGAGAAATCCAGATCCTGTGGACGCCCTTATTTATTTAG 1020
QY 1296 AGGGAATCCAGTGTAGAGTGGAGTACTGACACTGACAAATGCTCAGACGAGAAAGG 1355
|||||
Db 1021 ATGGATCCCAATGAGGAGGAGTACTGACAACTGACAAATGCTCAGACGAGAAAGG 1080
QY 1356 ACTGCGGTGCGCTCCAACTATTACCCGATTCGAAGCCTAGAGGCTCTTCTGAAACA 1415
|||||
Db 1081 ACTGCGGTGCGCTCCGAAATGTCACCCGGTTCCAAAGCCTAGAGGCTCTTCTGAGCAA 1140
QY 1416 GGCACAACTGACAAAGGCTGGGGTGGAGAGTGCATCCACAGAAATGAGCAGAACTAT 1475
|||||
Db 1141 GGCACAACTGACAAAGGCTGGGGTGGAGAGTGCATCCACAGAAATGAGCAGAACTAT 1200
QY 1476 CAAGGACATCTTCAATTAAGTGCAGAGAAACCTGCCAAGCCTTGTCATCTATGACA 1535
|||||
Db 1201 CGAGGACATCTTCAATTAAGTGCAGAGAAACCTGCCAAGCCTTGTCATCTATGACA 1260
QY 1536 CCACACTGCGATAGTGGACCCGACGATCTACCCAAATGCTGGCTTATCAAGACTAC 1595
|||||
Db 1261 CCACACTGCGATAGTGGACCCGACGATCTACCCAAATGCTGGCTTATCAAGACTAC 1320
QY 1596 TGGCGAAATCCAGATCCTGTGGACGCCCTTGGTGTATACAAAGATCCCACTGTCAAG 1655
|||||
Db 1321 TGGCGAAATCCAGATCCTGTGGACGCCCTTGGTGTATACAAAGATCCCACTGTCAAG 1380
QY 1656 TGGGAGTACTGCAACCTGACAGAGATGCTCAGATCAGAAATGAGTGCCTTCCCTCCG 1715
|||||
Db 1381 TGGGAGTACTGCAACCTGACAGAGATGCTCAGATCAGAAATGAGTGCCTTCCCTCCG 1440
QY 1716 AATGATATCTGGCTCCAAAGCCTTAGAGCCTTTTGAACAACACTGAGGAAAC 1775
|||||
Db 1441 AATGATATCTGGCTCCAAAGCCTTAGAGCCTTTTGAACAACACTGAGGAAAC 1500
QY 1776 CCCGGGTACAGAGTCTACTACCAATTAAGGACAGATTACGAGGACATCTCCACC 1835
|||||
Db 1501 CCCGGGTACAGAGTCTACTACCAATTAAGGACAGATTACGAGGACATCTCCACC 1560
QY 1836 ACTGTCACAGAAAGTCTGCAAGCTTGTGATCTATGACACACACAGCTAGTACG 1895
|||||
Db 1561 ACTGTCACAGAAAGTCTGCAAGCTTGTGATCTATGACACACACAGCTAGTACG 1620
QY 1896 ACCCCAGAAATACCCAAATGCTGGCTGACACGAACTACTGCAAGAACTCAGATGCT 1955
|||||
Db 1621 ACCCCAGAAATACCCAAATGCTGGCTGACACGAACTACTGCAAGAACTCAGATGCT 1680
QY 1956 GAGATTCGCCCTTGGTGTAGACAGTGAATCCAGTGTCAAGTGGAGTACTGCAACCTG 2015
|||||
Db 1681 GAGATTCGCCCTTGGTGTAGACAGTGAATCCAGTGTCAAGTGGAGTACTGCAACCTG 1740
QY 2016 ACGCATGCGCTGAGACAGATCAAGTGTCTTGAACACTGCAAGTGGTCCCAAGTCCA 2075
|||||
Db 1741 ACGCATGCGCTGAGACAGATCAAGTGTCTTGAACACTGCAAGTGGTCCCAAGTCCA 1800
QY 2076 AGCAGAGAGCTTCTTGAAGAAAGACACAGAGCAAGGAGCCGAGGCTCCAGAAATG 2135
|||||
Db 1801 AGCAGAGAGCTTCTTGAAGAAAGACACAGAGCAAGGAGCCGAGGCTCCAGAGCTCC 1860
QY 2136 TACCATGATGATGACAGATTATGAGGCTCATCTTCAACCTGTCACAGAAAGACA 2195
|||||
Db 1861 TACCATGATGATGACAGATTATGAGGCTCATCTTCAACCTGTCACAGAAAGACA 1920
QY 2196 TGTCACTTGTGCTCTATGACACACACTGCGATCAGAGGACAAAGAAATTTATCCA 2255
|||||
Db 1921 TGTCACTTGTGCTCTATGACACACACTGCGATCAGAGGACAAAGAAATTTATCCA 1980
QY 2256 AATGATGGCTGACAGAGAACTACTGCGAGAACTCAAGATCTGAGATTTAGTCTTGGTGT 2315

Db 1981 GATGATGGCTGACAGAGAACTACTGACAGAACTCAAGATCTGAAATGCTCCCTTGGTGT 2040
|||||
QY 2316 TATACCATGATGCCCAATGATGAGATGATGCAACTGACACAAATGTCAGTACA 2375
|||||
Db 2041 TATACCATGATGCCCAATGATGAGATGATGCAACTGACACAAATGTCAGTACA 2100
QY 2376 GAATCAAGTGTCTGACAGCTCCAGCTGTTTCTGAAACAGCAGCAAGGAGCAAGC 2435
|||||
Db 2101 GAATCAAGTGTCTGACAGCTCCAGCTGTTTCTGAAACAGCAGCAAGGAGCAAGC 2160
QY 2436 CCCAGTCCAGAGTGTACTACATGATGATGACAGAGTATGAGGCTCATTTCCACC 2495
|||||
Db 2161 CCCGGGTGACAGAACTGCTACCATGATGATGAGAGAGTATGAGAGTTTCAATCTCCACC 2220
QY 2496 ACTGTACAGAGAAAGCATGTCAGTCTTGGTCTCTGACACACAGCTGAGGATCAGAGA 2555
|||||
Db 2221 ACTGTACAGAGAAAGTATGTCAGTCTTGGTCTCTGACACACAGCTGAGGATCAGAG 2280
QY 2556 ACCACGAATCTACCCAAATGTCCTGACAGAGAACTACTGACAGAAATCCAGATGCT 2615
|||||
Db 2281 ACCATGAATCTACCCAAATGTCCTGACAGAGAACTACTGACAGAAATCCAGATGCT 2340
QY 2616 GAGATTCGCCCTTGGTGTATACATGATCCAGTGTCCAGATGAGTACTGCAACCTG 2675
|||||
Db 2341 GAGATTCGCCCTTGGTGTATACATGATCCAGATGATCCAGATGAGTACTGCAACCTG 2400
QY 2676 AGCAATGTCCAGTATGGAATCAACTCTCTCAAACTCCAGGCTGCCATGCTCA 2735
|||||
Db 2401 ACCAATGTGTGTATGGAATCAAGTGTCTTCAAACTCCAGGCTGCCATGCTCA 2460
QY 2736 AGCAGAGACTCTCTCTGAAAGAACACAACTGAAACAGCACTGGGTCCAGAGACTGC 2795
|||||
Db 2461 AGCAGAGAGTCTCTCTGAAAGAACACAACTGAAACAGCACTGGGTCCAGAGACTGC 2520
QY 2796 TACCGAGTATGAGACAGAGTATTCAGAGGACACTGCCACATCTACAGAGAAACA 2855
|||||
Db 2521 TACCAAGTATGAGACAGAGTATTCAGAGGACACTGCCACATCTACAGAGAAACA 2580
QY 2856 TGTCACTTGTGGTGTCTATGACACACAAATGTCATGAGATGCCATTTATCTATCCA 2915
|||||
Db 2581 TGTCACTTGTGGTGTCTATGACACACAAATGTCATGAGATGCCATTTATCTATCCA 2640
QY 2916 AATGTCGCTGACAGAGAACTACTGACAGAACTCAGATCTGAGATTCGCTTGGTGT 2975
|||||
Db 2641 AATGTCGCTGACAGAGAACTACTGACAGAAATGAGATCTGAGATTCGCTTGGTGT 2700
QY 2976 TACACCATGATCCCAAGTGTAGAGTGGAGTACTGCAACTGACAGATGTCAGTACA 3035
|||||
Db 2701 TACACCATGATCCCAAGTGTAGAGTGGAGTACTGCAACTGACAGATGTCAGTACA 2760
QY 3036 GAATCAGATGTCCTCAACAATCCCAAGTGGCCCGGTTCCAAAGCAGAGAGCTCCTCT 3095
|||||
Db 2761 GAATCAGATGTCCTCAACAATCCCAAGTGGTCCCGGTTCCAAAGCAGAGAGCTCCTCT 2820
QY 3096 GAACAAGCAGACCTGAGAAAGGCTGTGGTCCAGATGTCACAGTGTATGAGAGG 3155
|||||
Db 2821 GAACAAGCAGACCTGAGAAAGGCTGTGGTCCAGATGTCACAGTGTATGAGAGG 2880
QY 3156 AGTTATCGAGCATTCCTCCACAGTGTCAAGAGAACTGTCATCTTGTGATCT 3215
|||||
Db 2881 AGTTATCGAGCATTCCTCCACAGTGTCAAGAGAACTGTCAGTCTTGTGATCT 2940
QY 3216 ATGATTCACAGTGTGATCAGAGGACCCGAAAGCTACCAATGCTGGCTGACGAG 3275
|||||
Db 2941 ATGATTCACAGTGTGATCAGAGGACCCGAAAGCTACCAATGCTGGCTGACGAG 3000
QY 3276 AACTACTGAGGAATCCAGATTCGGAAGAACACCTGTGTATCAACAACGATCCGTGT 3335
|||||
Db 3001 AACTACTGAGGAATCCAGATTCGGAAGAACACCCCTGTGTATCAACAACGATCCAGT 3060
QY 3336 GTGAGTGGAGTACTGGAATCTGACAAATGCTCAGAAACAAATCAGATGCTCTAGAG 3395
|||||

Db 3061 GTGAGGTGGAGTACTGCAACCTGACAAATGCTCAGAAAAGAAATGAGTGTCTAGAG 3120
QY 3396 ACTCCACAGTGTGTTCAGAGTTCGAAGCATGAGAGCTATTCTGAGACAGACCAACTGAG 3455
Db 3121 ACTCCACAGTGTGTTCAGAGTTCGAAGCATGAGAGCTATTCTGAGACAGACCAACTGAG 3180
QY 3456 CAAACCCCTGTGTGTCGGCAGTGTCACTGATGATGAGCAGAGTATTCAGAGCAGATTTC 3515
Db 3181 CAAACCCCTGTGTGTCAGAGAGTGTCACTGATGATGAGCAGAGTATTCAGAGCAGATTTC 3240
QY 3516 TCACACAGTGTGACAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 3575
Db 3241 TCACACAGTGTGACAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
QY 3576 CAGAGACCCCGAGAAATACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 3635
Db 3301 AAGAGACCCCGAGAAATACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
QY 3636 GATGCCGATGACAGAGCCCTGTGTGTTCATGATGATGATGATGATGATGATGATGATGAT 3695
Db 3361 GATGCCGATGACAGAGCCCTGTGTGTTCATGATGATGATGATGATGATGATGATGATGAT 3420
QY 3696 AACCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3755
Db 3421 AACCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
QY 3756 GTTCCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3815
Db 3481 GTTCCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3510
QY 3816 CGGAGGAG 3875
Db 3511 ----- 3510
QY 3876 CCCCATGACAGACAGACGTTTCATTCAGAGGACCAATAATGAGGAGTCTGGAAAAAT 3935
Db 3511 ----- 3510
QY 3936 TACTGCCGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3995
Db 3511 ----- 3510
QY 3996 AAACCTTTTGACTACTGATATATCCCTCTGTGTGATGATGATGATGATGATGATGATGAT 4055
Db 3511 -----GCATCCTTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 3537
QY 4056 CCTCAAGTGGAGCCGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4115
Db 3538 CCTCAAGTGGAGCCGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3597
QY 4116 CATTCCTGGCCCTGGCAAGTCAAGTCAAGACAGAGTTGGAAAGCACTTCTGTGAGAGC 4175
Db 3598 CATTCCTGGCCCTGGCAAGTCAAGTCAAGACAGAGTTGGAAAGCACTTCTGTGAGAGC 3657
QY 4176 ACCTTAATATCCGACAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 4235
Db 3658 ACCTTAATATCCGACAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3717
QY 4236 CCTTCATCCTCAAGAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4295
Db 3718 CCTTCATCCTCAAGAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3777
QY 4296 CAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4355
Db 3778 CAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3837
QY 4356 AAGCTAAGCAGGCTGCTGCTATCACTGACAAAGTATGATGATGATGATGATGATGATGAT 4415
Db 3838 AAGCTAAGCAGGCTGCTGCTATCACTGACAAAGTATGATGATGATGATGATGATGATGAT 3897
QY 4416 GACTACATGCTGACGCGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4475
Db 3898 AATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3957

QY 4476 ACCTTTGGAGCTGCTGCTCAAGAGAGCCAGCTCTGTTATGATGATGATGATGATGATGATGAT 4535
Db 3958 ACCTTTGGAGCTGCTGCTCAAGAGAGCCAGCTCTGTTATGATGATGATGATGATGATGATGAT 4017
QY 4536 AATCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4568
Db 4018 AATCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4077
QY 4569 GCCAGAGCAGTACAGTGTGAG 4628
Db 4078 GCCAGAGCAGTACAGAGATGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4137
QY 4629 GACAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4688
Db 4138 GACAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4197
QY 4689 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4748
Db 4198 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4257
QY 4749 AATTAATGAGCGGAG 4808
Db 4258 AATTAATGAG 4317
QY 4809 AGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4868
Db 4318 GGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4377
QY 4869 CAGCTATGCCAAACCTTGGCATTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 4928
Db 4378 CAGCTATGCCAAACCTTGGCATTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 4437
QY 4929 GACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4988
Db 4438 GCTGGTCTTCCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4497
QY 4989 AATTTGATTTGA 5000
Db 4498 AATTTGATTTGA 4509

RESULT 3
145623
LOCUS 145623 2753 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5637492.
ACCESSION I45623
VERSION I45623.1 GI:2469725
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2753)
AUTHORS Dawson,K., Edwards,R.M. and Forman,J.M.
TITLE Activatable fibrinolytic and anti-thrombotic proteins
JOURNLS Patent: US 5637492-A 1 10-JUN-1997;
FEATURES
source location/Qualifiers
1..2753
BASE COUNT 766 a 672 c 692 g 623 t
ORIGIN

Query Match 24.7%; Score 1237; DB 6; Length 2753;
Best Local Similarity 81.0%; Pred. No. 0;
Matches 1512; Conservative 0; Mismatches 275; Indels 80; Gaps 3;

QY 3159 TATGAGGATATTCCTGACACATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3218
Db 911 TATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 970
QY 3219 ATACACACTGCGATCAG 3278
Db 971 ACCCTTACACACATTAACAG 1030

QY	3279	TACTGCAGATCCAGATTCTGGGGAAACAACCCCTGGTATTACAAACCCATCCGTGTG	3338
Db	1031	TACTGCCGCAATCCGACGGAAAAAGGCCCCCATGGTCCCTACCAACCAAGCCAAAGG	1090
QY	3339	AGCTGGAGTACTGCATCTGCACACATCCTCAGAAACGAAATCAGGTGTCTAGAGACT	3398
Db	1091	CGGTGGAGTACTGTAAGATTAACGTCCTG-----	1159
QY	3399	CCACAGCTGTTCCAGTTCACACATAGSAGGCTCAATCTGAGAGACCAACTGAGCA	3458
Db	1120	-----TGACTCTCCCAAGTATCCACAGCAACATTGGCTCCACACAGCACCTGAGCTA	1174
QY	3459	ACCCCTGTGTCGGCAGTGTCACTACATGTAAATGAGCCAGATTATCGAGGCACATTTCC	3518
Db	1175	ACCCCTGTGTCGAGACTGCTACATGTGTATGACAGACACTCCAGGACATCTCC	1234
QY	3519	ACCACGTGCAGAGAGAGATGTCAATTTGGTATCTCATGACACACACCGCATCTAG	3578
Db	1235	ACCACACACAGGAAAGAAAGTGTCACTGTGTATCTATGACACACACCGGACACAG	1294
QY	3579	AGGATCCCGAANAATACCCAAATGATGGCCCGACAATGAATCTACTCGAAGATTCAGT	3638
Db	1355	GCGGATTAAGGCGCCCTGGTGTATTACCAACAGACCCCAACGTCAGTGGGAGTACTCAAC	1414
QY	3699	CTACGCCGATGCTGCACACAGAAAGGACGTGGTGGCTCTCCGATGTGATCCAGGTT	3758
Db	1415	CTGAAAAAATGCTCGAAGAACAGGAGATGTTATAGCACTCCGCTGTGTGTCCTGCT	1474
QY	3759	CCAAAGCTTAGGCGCTCTTCTGAACAAGACTGATATGTTGGAAATGGAAAGATACGG	3818
Db	1475	CCAGATGTAGAACACTCCTCCGAAAGAACTGATATGTTGGAAATGGAAAGATACCGA	1534
QY	3819	GCGAAGAGGCAACACACTGTTACTAGGAGACCATCCACAGAAATGGGTCGCCAGAGCC	3878
Db	1535	GCGAAGAGGCGCACCTGTTACTGGGACGCGCATGCCAGAGACTGGCTGCCAGAGGCC	1594
QY	3879	CATAGACACAGACACTTCATCCAGGAGCAAAATTAATGGGAGGCTCTGGAAAAATTAC	3938
Db	1595	CATAGACACAGATTTTCTTACTCCAGAGACAAATTCAGCGGGGGGCTCTGGAAAAATTAC	1654
QY	3939	TGCGCTAACCCCTGATGTGACATCAATGTGTCCCTGGTCTCACAAATGAATCCAGAAAA	3998
Db	1655	TGCGCTAACCCCTGATGTGAGTGTAGGTGGTCCCTGGTCTCACAGCAAAATCCAGAAAA	1714
QY	3999	CTTTTGTACTGTGATATCCCTCTGTGGATCCTCTCATATTGATGTGGGAAGCCT	4058
Db	1715	CTTTTACGACTACTGTGATGTCTCTCACTGTGGGCCCTTATTTTGAATGTGGGAAGCCT	1774
QY	4059	CAAGTGAGCCGAGAAAAATGTCCTGGAAGCATTTGAGGGGGTGTGTGCCACCCACAT	4118
Db	1775	CAAGTGAGCCGAGAAAAATGTCCGGAAGGCTGTGAGGGGGTGTGTGCCACCCACAT	1834
QY	4119	TTCGTGCCCTGTGCAGTCAAGTCTCAGAACAAAGTTTGGAAAGCACTCTGTGGAGCAC	4178
Db	1835	TTCGTGCCCTGTGCAGTCAAGCTTAAACAAAGTTTGGAAATGCACTCTGTGGAGCAC	1894
QY	4179	TTTAATATCCCAAGAGTGGGTGTGACTGCTCACTGCTTGAAGAAAGTCTTAAGGCT	4238
Db	1895	TTTAATATCCCAAGAGTGGGTGTGACTGCTGCCCCTCTTGGAGAAAGTCTCCCAAGGCT	1954
QY	4239	TCAATCTACAAGGTCATCTGTGGGACACACAAGAAAGTGAATCTCGAATCTCAATGTTCAG	4298
Db	1955	TCAATCTACAAGGTCATCTGTGGGTGACACACAAGAAAGTGAATCTCGAACCGCATGTTCAG	2014
QY	4299	GAATATGAAGTGTCTAGGCTGTCTTGTGAGCCCAACAAGCAGATATTGCTTGTGAAG	4358
Db	2015	GAATATGAAGTGTCTAGGCTGTCTTGTGGAGCCCAACAAGAAAGTATTGTGCTTGTGAAG	2074

OY	4359	CTGGCAGGCGCTGCCTCAATCTGCACAAAGAATAGCCAGGTTCCTTCCCATCCGAC	4418
Db	2075	CTAAGCAGTCTCTGCCGTACTCACACAAAGTAATCCAGTTGTCTGCATCCCAAAT	2134
OY	4419	TACATGGTCACGCCACGAGCACTGAATGTTATCATCTACTGCTGGGAGAAAACCAGGTACC	4478
Db	2135	TATGTGTGCTGTACCGGACGAAATGTTTCACTGGCTGGGGAGAAACCCAAGTACT	2194
OY	4479	TTTGGAAGCTGGCCCTTCTCAAGGAAAGCCGACGCTCTTGTATTATGAAATGAATGTGCANT	4538
Db	2195	TTTGAGACTGGCCTTCTCAAGGAAAGCCAGCTCCCTGTGATTGGAATTAAGTGTCAAT	2254
OY	4539	CACATATAAGT-----AATTTGGCTGCAATTTGGCC	4571
Db	2255	GCGTATAGATTTTCTGAATGGAAAGATCCAAATCCACCGAACCTCTGTGGGCAATTTGGCC	2314
OY	4572	AGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGGCCTCTGGTTTCTTCCGAAAGGAC	4631
Db	2315	GGAGGCACTGACAGTTGCCAGGGTGACAGTGGAGGCTCTGTGGTTTCTTCCGAAAGGAC	2374
OY	4632	AAATACATTTTACAAGAGTCACTTCTTGGGGTCTTGCTGTGCAAGCCCCCAATTAAGCT	4691
Db	2375	AAATACATTTTACAAGAGTCACTTCTTGGGGTCTTGCTGTGCAAGCCCCCAATTAAGCT	2434
OY	4692	GGTGTCAATCTGTGTTTTTCAAGTTTGTACTTGGATTGGGGAATGATGAATAATAT	4751
Db	2435	GGTGTCAATCTGTGTTTTTCAAGTTTGTACTTGGATTGGGGAATGATGAATAATAT	2494
OY	4752	TAAATGACGGGAGACAGAGTAGAGTCAACACTTCTTAGAGCTGAACGTTGGGTAAAG	4811
Db	2495	TAAATGACGGGAGAGAGTAGAGTCAACACTTCTTAGAGGCTGGAACGTTGGGTAAAG	2554
OY	4812	AATTAGCATCTGGAATATATAGACAGCAATCAACGAAGACACTGTTCCACACTCCAG	4871
Db	2555	AATTAGCATCTGGAATATATAGACAGCAATCAACGAAGACACTGTTCCACACTCCAG	2614
OY	4872	CTAGCCCAACCTGGCATTTTGGATTTTGGATTTATTAAGCTTTAAGGTCGAGTGCAC	4931
Db	2615	CTAGCCCAACCTGGCATTTTGGATTTTGGATTTATTAAGCTTTAAGGTCGAGTGCAC	2657
OY	4932	AAATTCGTATTAAGGTGTATAGCTATGATGACATTTGTTAAAAATTAACCTGCACATTAT	4991
Db	2658	GGATTCGTGTAGTAGGATGACATAGCATGACATTTGTTAAAAATTAACCTGTACTTAAC	2717
OY	4992	TTGATTT 4998 Db 2718 TTTGATTT 2724	
RESULT 4			
HSPMR			
LOCUS	HSPMR	2732 bp	linear PRI 21-MAR-1995
DEFINITION	Human mRNA for plasminogen.		
ACCESSION	X05199		
VERSION	X05199.1 GI:35530		
KEYWORDS	plasminogen; signal peptide.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	I (bases 1 to 2732)		
JOURNAL	Forsegren,M., Raden,B., Israelsson,M., Larsson,K. and Heden,L.O.		
MEDLINE	Molecular cloning and characterization of a full-length cDNA clone		
PUBMED	FEBS Lett. 213 (2), 254-260 (1987)		
COMMENT	87162490		
FEATURES	3030813		
Source	Data kindly reviewed (21-DEC-1987) by HEDEN L.-O. Location/Qualifiers 1..2732 /organism="Homo sapiens" /db_xref="taxon:9606"		
CDS	55...2487		


```
/note="precursor polypeptide (AA -19 to 791)"
/codon_start=-1
/protein_id="CAA28831.1"
/db_xref="GI:35531"
/db_xref="SWISS-PROT:P00747"
/translation="MEKHEVILLILLFLKSGGPELDYVNTNGASLSEPVNKKOLGAG
SIEEACAKCEDEEPPCRAPQVSHKROCKVIAENRKSIIIRMDVYLPERKYLISE
CTKGNKNTYGTSTKRNKITCCKMSSTPHRPRSPATHSEGLEENECNPNDRQ
GPMVCTTDPKRYDCDILSEBECMHGSEVTDKISKTHSGLECOQWDSQSHAG
YIPSKFPNKLKKNYCRNDRRLRWCFTTIDNKKWELCDIPRCTPPSSGPTQCL
KGTENYRGNAVAVTSGHSCQHSQAOTPHTRNREBENPCNKLDENTCRNDRGRAM
CHTNGSVMREYCKIPSCDSPEVSTQLAPLAPPELTIVYDQCYGDCQSTRGSTT
TTGKCSQSMSPHRRHOKTPEYRNAGLTMVCRNPADKGLTDLKLSPPAVITDKYI
LKSGTEASVMAPPVYLLPVEYTPSEEDCKMGNGKGRKGRATVYGTCCODMAQO
EPHRSYTPETNPRAKLEKNCRNDRPQVCKYTPNPKRKYCDVPOCAAQSPD
CGKPOVEPKCPGRVYVGGCYAHPHSPWQVSLRTFRGHFECGTLISEBYVLTAHCL
EKSPRSSTYVILGAHOEVNLEPHVDELSPVRLKIDLPKIDIALKLSPPAVITDKYI
PACLPSNYVYADRTCEFTTGWGTEGTGAGLLEADLPVLENKVRNRYEFLNGRVO
STELCAGHLAGGTDSCQGDGSGPLVCFEKDYILQGVTSWGLGCAHPNKPVYVRSR
FVWIEGVMMNN"
55..111
112..2484
/product="put. mature plasminogen (AA 1-791)"
misc_feature
1792..1797
/note="cleavage site"
misc_feature
2689..2694
/note="pot. polyA signal"
polyA_site
2732
/note="polyA site"
BASE COUNT 758 a 669 c 688 g 617 t
ORIGIN
Query Match 24.7%: Score 1235.4; DB 9; Length 2732;
Best Local Similarity 80.9%: Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 276; Indels 80; Gaps 3;
```

```

Db 1405 CTGAAAAAATGCTCAGAGAACGAGACGAGTGTATGACCTCGCCCTGTGCTGCTT 1464
Qy 3759 CCAGAGCTGGGCGCTCTTCTGAAACAGACTGTATGTTGGGAATGGGAAGGATACCGG 3818
Db 1465 CCAGATGTAGAGACTCTCCGAGAAAGACTGTATGTTGGGAATGGGAAGGATACCGA 1524
Qy 3819 GCGAAGAAAGCAACACACTGTACTGTGGACGCCATGCCAGAAAGGCTGCGCCAGAGGCC 3878
Db 1525 GCGAAGAGGCGACACACTGTACTGTGGACGCCATGCCAGAAAGGCTGCGCCAGAGGCC 1584
Qy 3879 CATAAGACACAGACCTTGATCTCCAGGAGCAAAATAATGCGAGGTCTGAAAAAAATATAC 3938
Db 1585 CATAGACACAGACTTGTACCTCAGAGCAAAATCCAGGGGGGGGTCTGAAAAAAATATAC 1644
Qy 3939 TGGCGTAAACCTTGATGTGACATCAATGATGCCCGGTGCTTACCAATGATCCAAAGAAA 3998
Db 1645 TGCCGTAAACCTTGATGTGATGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1704
Qy 3999 CTCTTTGACTACTGTGATATCCCTCTCTGTCATCTCTTCACTTGTATGTTGGAGGCTT 4058
Db 1705 CTCTTACGACTACTGTGATGTGATGCTCCATGTCGCGGCCCTTCATTTGATGTGGAGGCTT 1764
Qy 4059 CAGGTGAGCCGCAAGAAATGCTCTGGAAGCAATGTAGGGGGGTGTGGCCACCACAT 4118
Db 1765 CAGGTGAGCCGCAAGAAATGCTCTGGAAGGCTGTGGGGGGGTGTGGCCACCACAT 1824
Qy 4119 TCCTGGCCCTGGCAAGTCAAGTCTCAGAACCAAGTTTGGAAAGCACTCTGTGGAGGCAAC 4178
Db 1825 TCCTGGCCCTGGCAAGTCAAGTCTTAGAACCAAGTTTGGAAATGCACTCTGTGGAGGCAAC 1884
Qy 4179 TTAATATCCCAAGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4238
Db 1885 TTGATATCCCAAGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
Qy 4239 TCATCTCAAGAGTCAATCTCTGGGTGCACACCAAGAGTGAACCTCGAATCTCATGTTCCAG 4298
Db 1945 TCATCTCAAGAGTCAATCTCTGGGTGCACACCAAGAGTGAACCTCGAATCTCATGTTCCAG 2004
Qy 4299 GAAATGAAGTGTCTGAGGTGTTCTTGGAGGCCACACAGACGATATGTCCTTGCATTAAG 4358
Db 2005 GAAATGAAGTGTCTGAGGTGTTCTTGGAGGCCACACAGAAAGATATGTCCTTGCATTAAG 2064
Qy 4359 CTAAAGAGGCTGCGCTCATCTAGACAAAGATATGCCAGCTGTGCTGCCATCCAGAC 4418
Db 2065 CTAAAGAGTGTGCGCTCATCTAGACAAAGATATGCCAGCTGTGCTGCCATCCAGAAAT 2124
Qy 4419 TACATGTCACGCCGACGACGATGTTACATCACTGCTGGGAGAGAACCCAGAGTACC 4478
Db 2125 TATGTGTGCTGCTGACCGAGCCGAATGTTTCACTGCTGTGGGAGAGAACCCAGAGTACC 2184
Qy 4479 TTTGGGACTGGCCCTTCTCAAGAGACCCAGCTCCTGTATATGAAATGAAGTGTCCAT 4538
Db 2185 TTTGGGACTGGCCCTTCTCAAGAGACCCAGCTCCTGTGATTTGAAATGAAGTGTCCAT 2244
Qy 4539 CACTATATAGT-----ATATTTGCTGAGCAATTTGGCC 4571
Db 2245 CGGTATGATGTTTGTGAATGGAAGATCCAAATCCAGCACTGTGCTGGGCAATTTGGCC 2304
Qy 4572 AGAGGCACTGACAGTTGCCAGAGGTGACAGTGAAGGCTCTGTGTTTCTCTGAGAGAGAC 4631
Db 2305 GGAGGCACTGACAGTTGCCAGAGGTGACAGTGAAGGCTCTGTGTTTCTCTGAGAGAGAC 2364
Qy 4632 AAATATCATTTTACAAAGAGTCACTTCTTGAGGTCTTGAGCTGTGACGCCCCCAATAAGCCT 4691
Db 2365 AAATATCATTTTACAAAGAGTCACTTCTTGAGGTCTTGAGCTGTGACGCCCCCAATAAGCCT 2424
Qy 4692 GGTGTCTAAGCTGTGTTCAAGGTTTGTATCTGTGATGAGGAAATGATGAATAATAT 4751
Db 2425 GGTGTCTAAGCTGTGTTCAAGGTTTGTATCTGTGATGAGGAAATGATGAATAATAT 2484
Qy 4752 TAAATTGACGGGAGACAGATGAAGCATCAACTATGAAAGCTGAAGCTGGGTAAG 4811
```

FEATURES	source
LOCUS	2753 bp
DEFINITION	plasmidogen CDNA.
ACCESSION	A22096
VERSION	A22096.1 GI:641445
KEYWORDS	plasmidogen.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.
FEATURES	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
LOCATION/Qualifiers	1..2753

CDS

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
65..2497
/codon_start=1
/product="plasmaogen"
/protein_id="CA01587.1"
/db_xref="GI:64146"
/translation="MEHKEVALLLLFLKSGGCEPLDDYVNTQGASLFSYTKOLGAG
STEEACACEDEEFTCRARQYHSKRQOCVITAMENKSSIIRMDVYLFEKYLGLS
CGTKGNKYNTRGMSKNTGICQKWSSTSPHREFSPTAHPEEGEELMNCSDPNDDP
GWCATYTPKERYDYCDLKEEBECMHGSEENDGKISTMGHGLCOAMDSCSPHAFH
YIYSPFNKLNKKNYCRNPDELRLPWCFTDDPKRHELDIIRCTYTPSSGAYOCL
KGTNGRKNGLAVYVSGHTQCHWSAOTPHHNTPEFNPKCLMDNYCRNDGRAPRA
CHTNSQYRMWEYCKIPSCDSPLYSTGLAFTAPPELTYPVOCCYCHGDSRGSSST
TTRKGSQSSMPHRRCKRQPNYRNAGLGNMCRPDADRKGRKATYVGTGTCOPMAAE
LAKKSGTEASVAPRPVYLLPYETSEBECMGNGKRGKATYVGTGTCOPMAAE
EPRHRIETPCPNFRAGLEKNYCRNPDGDGGWCATYPRKLYDVCVPOCAAPARCF
CGKRGYKPKCGKRVVGGCAVABHPSMWQVSLSTRFGMHFGCGTLISPEVYLAHIC
EKSPSSRYVILGAHDEVNLBPHGOEIEVSRLEFTEPRKDALTLKLSPEVATDKVY
PACPSRYVYVADRTCECFITMGETGTGAGALKEAOLVIEENKVCNRRYELNGRVO
STELCAGLILAGGTDSGCGSDGSPLYVCFEKDKYTLIGVTSMLGICARPNKPGYVVRSH
FVTYMEGVRNRN"
```

BASE COUNT	766 a	672 c	693 g	622 t
ORIGIN				
Query Match	24.7%	Score 1235.4;	DB 6;	length 2753;
Best Local Similarity	80.9%	Pred. NO. 0;		
Matches 1511; Conservative	0;	Mismatches 276;	Indels 80;	Gaps 3;

QY	3159	TATCGAGGCATATTCCCTCCACCACTGTCACAGAAAGACCTGTCAATCTTGSCATCATTANG	3218
Db	911	TATCGGGGGAATGTGGCTGTTACGCTGTCCGGGACACTGTCTCAGCACTGGATGTACAG	970
QY	3219	ATACCACTGGCATCAGAGACCACAGAAAACTATCCCAATGCTGGCTGACCGAGAAC	3278
Db	971	ACCCCTCACACATATACAGAGACACAGAAAATTCTCCGTGCAAAATTTGGATGAAAC	1030
QY	3279	TACTGGAGAATCCAGATTCTGGGAAACACCTGTGTTACACAACCATCCGTGTGTCG	3338
Db	1031	TACTGGCAGAAATCTGACAGAAAAAGGCCCATATGTGCATATCAACCAACCGCAAAGT	1090

[illegible]

Db	2135	TATGTGCTGCTGACCGGACCGAATGTTTCACTGCTGGGGAAACCCAAAGGTA	CT	2194
QY	4479	TTTGGACATGAGCCCTTCACAGGAAGCCAGCTCCCTTGTATTTAGAAATGAAGTGTGCAT		4538
Db	2195	TTTGGAGCTGGCCCTTCACAGGAAGCCAGCTCCCTGTATTTAGAAATGAAGTGTGCAT		2254
QY	4539	CACATTAAGT-----	ATATTGTGCTGAGATTTGGCC	4571
Db	2255	CGCTATGAGTTCCTGTAATGGAAGAGTCCATCCACCGAAGCTGTGCTGGCAATTTTGGCC		2314
QY	4572	AGAGGCACTGACAGATTTGCCAGGCTGACAGTGGAGAGGCTCTGGTTGCTTCGAGAAGAC		4631
Db	2315	GGAGGCACTGACAGATTTGCCAGGCTGACAGTGGAGAGGCTCTGGTTGCTTCGAGAAGAC		2374
QY	4632	AAATACATTTTCAAGAGATCCTCTGGGGGCTTGGCTGTGTCAGAGCCCAATTAAGCT		4691
Db	2375	AAATACATTTTCAAGAGATCCTCTGGGGGCTTGGCTGTGTCAGAGCCCAATTAAGCT		2434
QY	4692	GGTGTCTATGCTGTGTTCAAAGTTTGGTTACTTGGATTGAGGGAAATGATGAGAAATAT		4751
Db	2435	GGGTCTATGCTGTGTTCAAAGTTTGGTTACTTGGATTGAGGGAAATGATGAGAAATAT		2494
QY	4752	TAAATGAGGGGAGACAGATGTAACATCAACCTTAAGCTTAAGCTGAAGAGGGAAG		4811
Db	2495	TAAATGAGGGGAGACAGATGTAACATCAACCTTAAGCTTAAGCTGAAGAGGGAAG		2554
QY	4812	ATTAGCATGCTGGAAATATATAGACAGCAATCAAGAGACACTGTTCCACAGTACAG		4871
Db	2555	ATTAGCATGCTGGAAATATATAGACAGCAATCAAGAGACACTGTTCCACAGTACAG		2614
QY	4872	CTATGCAACACCTTGGCATTTTGGTATTTTGGTATTAAGCTTGAAGTCTGACTGAC		4931
Db	2615	CTATGCAACACCTTGGCATTTTGGTATTTTGGTATTAAGCTTGAAGTCTGACTGAC		2657
QY	4932	AAATCTGATTAAGTGTGATAGCATATTTGTTAAATTAACCTGCACTTAT		4991
Db	2658	GGATCTGATTAAGTGTGATAGCATATTTGTTAAATTAACCTGCACTTAT		2717
QY	4992	TTGATTT		4998
Db	2718	TTTGATTT		2724
RESULT 6				
HUMPMGM				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PubMed				
FEATURES				
Source				
gene				
mRNA				
CDS				

[illegible]

Db 694 CGGGGGGCTGAAAAAAATTAAGTCCGTAACCTGATGCTAGTGTAGTGTGCTCCG 753
Qy 3975 TGCTACACAAATGATCCAAAGAAATTTTGTACTGATGATATCCCTCTGTCATCC 4034
Db 754 TGCTACACAGCAAAATCCAAAGAAATTTTGTACTGATGATATCCCTCTGTCATCC 813
Qy 4035 TCTTCATTTGATTTGGGAAGCTCAAGTGGAGCCGAGAAATGTCCTGGAAGATTTGA 4094
Db 814 CCTTCATTTGATTTGGGAAGCTCAAGTGGAGCCGAGAAATTTGTCCTGGAAGGTTGTA 873
Qy 4095 GGGGGGCTGTGGCCCAACATTCCTGGCCCTGGCAGTGCAGTCTGAGCAAGAGTTT 4154
Db 874 GGGGGGCTGTGGCCCAACATTCCTGGCCCTGGCAGTGCAGTCTGAGCAAGAGTTT 933
Qy 4155 GGAAGACATTTCTGTGGAGGACCTTAATATCCAGAGTGGCTGATGCTGCTGAC 4214
Db 934 GGAATGACATTTCTGTGGAGGACCTTAATATCCAGAGTGGCTGATGCTGCTGAC 993
Qy 4215 TGCTGGAAGAAAGTCTCAAGGCTTATCCTTACAGTATCTCTGGTGCACACAGAA 4274
Db 994 TGCTGGAAGAAAGTCTCAAGGCTTATCCTTACAGTATCTCTGGTGCACACAGAA 1053
Qy 4275 GTGAACCTCGAATCTCATGTTTCAGAAATAGAAGTGTCTAGGCTGTTCTTGAGCCACA 4334
Db 1054 GTGAATCTCGAACCAGCATGTTTCAGAAATAGAAGTGTCTAGGCTGTTCTTGAGCCACA 1113
Qy 4335 CAACGACATATGCTGCTGCTTAAGTAAAGAGGCTGCTGCTGATCTGATGACAAAGTATG 4394
Db 1114 CGAAAGATATGCTGCTGCTTAAGTAAAGAGGCTGCTGCTGATCTGATGACAAAGTATG 1173
Qy 4395 CCAGCTTGTCTGCATCCCGAGACTATGTCACGCGCAGCAGATGATGTTATACACT 4454
Db 1174 CCAGCTTGTCTGCATCCCGAGACTATGTCACGCGCAGCAGATGATGTTATACACT 1233
Qy 4455 GGGTGGGAGAAACCAAGTACTGTTGGAGTGGCTGCTTCAAGAGCCAGCTGCTT 4514
Db 1234 GGGTGGGAGAAACCAAGTACTGTTGGAGTGGCTGCTTCAAGAGCCAGCTGCTT 1293
Qy 4515 GTTATGAGATGAAGTGTCAATCACTATAGT----- 4548
Db 1294 GTTATGAGATGAAGTGTCAATCACTATAGT----- 4548
Qy 4549 -ATATTTGCTGAGATTTGGCCAGAGGACTGACAGTGGCAGGAGAGAGG 4607
Db 1354 GAACTCTGTCTGCTGCTTTGGCCAGGAGGACTGACAGTGGCAGGAGAGAGG 1413
Qy 4608 CTTCTGTTGCTTCCAGAGGACAAATCAATTTTACAAGAGTCACTTTGGGCTT 4667
Db 1414 CTTCTGTTGCTTCCAGAGGACAAATCAATTTTACAAGAGTCACTTTGGGCTT 1473
Qy 4668 GGGTGTCCAGCCCAATAGGCTGCTGCTATGCTGCTGCTTCAAGGTTGTTACTTG 4727
Db 1474 GGGTGTCCAGCCCAATAGGCTGCTGCTATGCTGCTGCTTCAAGGTTGTTACTTG 1533
Qy 4728 ATTGAGGGAATGATGAGAAATTAATTTGAGCGGAGAGACAGAGTCAATCAACTAC 4787
Db 1534 ATTGAGGGAATGATGAGAAATTAATTTGAGCGGAGAGACAGAGTCAATCAACTAC 1593
Qy 4788 TTGGAACCTGAAGCTGGTGAAGGATTTGATGCTGCTGGAATTAATAGACACATCAAC 4847
Db 1594 CTGAGGCTTGAAGCTGGTGAAGGATTTGATGCTGCTGGAATTAATAGACACATCAAC 1653
Qy 4848 GAAAGACATTTCCAGCTACAGCTATGCCAACTTTGGCATTTTGTGTGT 4907
Db 1654 GAAAGACATTTCCAGCTACAGCTATGCCAACTTTGTGTGTGT----- 1709
Qy 4908 ATAACTTTTAAGCTGACATCAATTTCTATTAATAGTGTATGACATTTG 4967
Db 1710 -----TTCTGACTGCTGATTTCTTAAAGTGAACATGACATTTG 1756
Qy 4968 TTAATAATTAATCTGCTACTATTTT 4993
Db 1757 TTAATAATTAATCTGCTACTATTT 1782

RESULT 7
106212
LOCUS 106212 2679 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 14 from Patent EP 031944.
ACCESSION 106212
VERSION 106212.1 GI:590227
KEYWORDS
SOURCE
ORGANISM unknown.
REFERENCE
AUTHORS 1 (bases 1 to 2679)
Mullinbill, E.R., Berkner, K.L., Foster, D.C., Kumar, A.A., Mackay, V.L.
and Parker, G.E.
TITLE Co-expression in eukaryotic cells
JOURNAL Patent: EP 031944-A2 14 14-JUN-1989;
FEATURES
source Location/Qualifiers
BASE COUNT 750 a 664 c 670 g 595 t
ORIGIN
Query Match 24.6% Score 1228.8 DB 6; Length 2679;
Best Local Similarity 80.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 277; Indels 80; Gaps 3;
Qy 3159 TATCGAGGCATATCTCCACCATCTGTCACAGAGAGACCTGTCATCTGATCTATG 3218
Db 868 TATCGAGGCATATCTCCACCATCTGTCACAGAGAGACCTGTCATCTGATCTATG 927
Qy 3219 ATACACACCTGATCAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3278
Db 928 ACCCTCAGACACATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 987
Qy 3279 TACTGAGGAATCCAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3338
Db 988 TACTGAGGAATCCAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1047
Qy 3339 AGTGGAGTATCTGATCTGACATGTCATGCAAGTCTGCAAGAAACAGATTCAGAGACT 3398
Db 1048 CGGTGGAGTATCTGATCTGACATGTCATGCAAGTCTGCAAGAAACAGATTCAGAGACT 1076
Qy 3399 CCCACTGTTGTCAGATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3458
Db 1077 -----TGACTCTCTCCAGATATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1131
Qy 3459 ACCCTGTGTCCGCGAGTGTACATGATGATGATGATGATGATGATGATGATGATGATG 3518
Db 1132 ACCCTGTGTCCGCGAGTGTACATGATGATGATGATGATGATGATGATGATGATGATG 1191
Qy 3519 ACCACTGTACAG 3578
Db 1192 ACCACTGTACAG 1251
Qy 3579 AGGAGCCAG 3638
Db 1252 AAGAGCCAG 1311
Qy 3639 GCCGATACAGGCTCTGTTGTTTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3698
Db 1312 GCCGATACAGGCTCTGTTGTTTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1371
Qy 3699 CTGAGCGGATCTGAGACAG 3758
Db 1372 CTGAG 1431
Qy 3759 CCAGAGCTAGGCTCTCTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3818
Db 1432 CCAGAGCTAGGAG 1491
Qy 3819 GGAAG 3878
Db 1491 GGAAG

```
Db 1492 GGCAGAGGGGACACACTGTTACTGAGACGCCATGCCAGATGGGGTCCCGAGGAGGCC 1551
QY 3879 CATAGACAGACAGCTTCATTCAGAGGACAATAATATGGCAGGTCTGNAAAAAATAC 3938
Db 1552 CATAGACAGACATTTCTACTCCAGAGCAATAATCCAGCGGGGTCTGAAAAAATAC 1611
QY 3939 TGGCGTAACCTGATGGATCATCATAGTCCCGTGGGTGATGACAAATGATCCAGAAAA 3998
Db 1612 TGGCGTAACCTGATGGATCATCATAGTCCCGTGGGTGATGACAAATGATCCAGAAAA 1671
QY 3999 CTTTTGACTACTGTGATATCCCTCTCTGTGATCTCTCTTCATTTGATTTGGGAAACCT 4058
Db 1672 CTTTACGACTACTGTGATATCCCTCTCTGTGATCTCTCTTCATTTGATTTGGGAAACCT 1731
QY 4059 CAAATGAGACCGAGAAATATGCTGGAAGCAATTTAGGGGGGTGTGGCCACCCACAT 4118
Db 1732 CAAATGAGACCGAGAAATATGCTGGAAGCAATTTAGGGGGGTGTGGCCACCCACAT 1791
QY 4119 TCTTGCCCTGGCAAGTCTCAGAACAAAGTTTGGAAAGCACTTCTGTGAGGACCC 4178
Db 1792 TCTTGCCCTGGCAAGTCTCAGAACAAAGTTTGGAAAGCACTTCTGTGAGGACCC 1851
QY 4179 TTAATATCCCAAGAGTGGGTGCTGACTGCTGCTGCTGCTGGAAGAGTCTTCAAGGCT 4238
Db 1852 TTAATATCCCAAGAGTGGGTGCTGACTGCTGCTGCTGCTGGAAGAGTCTTCAAGGCT 1911
QY 4239 TCAATCTACAAAGTCACTCTGGGTGACACCAAGTGAAGTCAATCTCATTTTCAG 4298
Db 1912 TCAATCTACAAAGTCACTCTGGGTGACACCAAGTGAAGTCAATCTCATTTTCAG 1971
QY 4299 GAATATGAAGTGTAGGCTGTCTTGGAGGCCACACAGAGATATGCTTCTGCTTAAG 4358
Db 1972 GAATATGAAGTGTAGGCTGTCTTGGAGGCCACACAGAGATATGCTTCTGCTTAAG 2031
QY 4359 CTAAGAGGCGTCCGCTCTCATGACAAAGTAATGCCAGCTCTGCGCATCCCAAGC 4418
Db 2032 CTAAGAGGCGTCCGCTCTCATGACAAAGTAATGCCAGCTCTGCGCATCCCAAGC 2091
QY 4419 TACATGTACACCGCCAGAGCAATATTCATCATCTGAGTGGGAGAGAACCCAGGTACC 4478
Db 2092 TATGTGTGTGCTGACCGAGCCAGATGTTTCATCACTGCTGGGAGAGAACCCAGGTACT 2151
QY 4479 TTTGGAGCGGCTCTCTCAAGAGACCCAGCTCCCTGTTATGATGATGAAGTGTGCAT 4538
Db 2152 TTTGGAGCGGCTCTCTCAAGAGACCCAGCTCCCTGTTATGATGATGAAGTGTGCAT 2211
QY 4539 CACTATTAAGT-----ATATTTGCTGTGACATTTGGCC 4571
Db 2212 CGCTATGAGTTTCTGAATGGAAGATCCATTCACACCGAACTCTGTGCTGGGCAATTTGGCC 2271
QY 4572 AGAGGACGTACAGTGGCCAGGCTGACAGTGGAGGCGCTGTGTTCTCTCGAAGAGAC 4631
Db 2272 GAGGACGTACAGTGGCCAGGCTGACAGTGGAGGCGCTGTGTTCTCTCGAAGAGAC 2331
QY 4632 AAATACATTTTACAGAGTCACTTCTGGGCTCTTGGCTGTGACGCCGCCCAATTAAGCCT 4691
Db 2332 AAATACATTTTACAGAGTCACTTCTGGGCTCTTGGCTGTGACGCCGCCCAATTAAGCCT 2391
QY 4692 GGGTCTATGCTGCTGTTCAGAGTTTGTACTTGATGATGAGAGAAATGAATTAAT 4751
Db 2392 GGGTCTATGCTGCTGTTCAGAGTTTGTACTTGATGATGAGAGAAATGAATTAAT 2451
QY 4752 TAAATGACGAGACAGATGAGATCACTTACTTGAAGAGCTGAAACGTGGGTAAAG 4811
Db 2452 TAAATGACGAGACAGATGAGATCACTTACTTGAAGAGCTGAAACGTGGGTAAAG 2511
QY 4812 ATTATGACGCTGGAATATATAGACAGATCAAGAAAGAAAGCACTTCCAGCTACACAG 4871
Db 2512 ATTATGACGCTGGAATATATAGACAGATCAAGAAAGAAAGCACTTCCAGCTACACAG 2571
QY 4872 CTATGACCAACCTTGGCAATTTTGTATTTTGTGTATTAAGCTTTAAAGTCTGACAGAC 4931
Db 2572 CTATGACCAACCTTGGCAATTTTGTGTATTTTGTGTATTTTGTGTATTTTGTGTATTT 2614
```

```
QY 4932 AAATTCGTATTAGGTGTCAATAGTACATTTGTTAAAAATTAACCTGTGACTTATT 4991
Db 2615 GGAATCTGTAGTAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2674
QY 4992 TT 4993
Db 2675 TT 2676

RESULT 8
MACPEPSG LOCUS 2711 bp mRNA linear PRI 04-AUG-1993
DEFINITION Rhesus monkey plasminogen mRNA, complete cds.
ACCESSION J04697.1 GI:342272
VERSION 304697.1 GI:342272
KEYWORDS plasminogen.
SOURCE Rhesus monkey liver, cDNA to mRNA.
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE Tomlinson,J.E., Mclean,J.W. and Lawn,R.M.
TITLE Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
AUTHORS synthesis
J Biol. Chem. 264 (10), 5957-5965 (1989)
JOURNAL 69174660
MEDLINE 2925643
COMMENT Draft entry and computer readable sequence for [1] kindly provided
by J.E.Tomlinson, 19-JAN-1989.
FEATURES
source
1..2711
location/Qualifiers
1..2711
/organism="Macaca mulatta"
/db_xref="taxon:9544"
<1..2711
/product="plasminogen mRNA"
/11..2483
/note="plasminogen"
/54..2486
/note="plasminogen precursor"
/codon_start=1
/protein_id="AA36901.1"
/db_xref="GI:342273"
/translation="MEHEKYVLLLLFLKSGEBPLDYNKRGASLFSITKQAG
SIEBCAKCEBEFEFTKRSFYHSKEDQCYHANKRSSIVRMKVLYFEKVTLSL
CTGNGKNTRGMTSKRTGTCQWSSSTSPHRTFSPHPSGLSEACNDNDGQ
GWCYTDPEERFDYCDIPCEDECMHCSBENDGKISKMSGLEQANDSCPHAG
YLPSEFPKNTLKNKRNPDGEPFPCFTDIPNKRMELCDIPCTTPPPSSGPTVQCL
KGTGENYRDVAIVTSGHGTGMSAORTPHTHNRPENFPCKNLDENYCRNPGEKAPW
CYTTSQVSWMEYCKIPSCSSPVSSTEDLPAPPELTPVVOECYHGDGOSYSTSTT
TYGKCCQSWSSMTPIHHEKTPENPMNGLMNTCRNPDADKGPWCCTPDSVMEYCN
LKCSGTBSVAAPPVPAQOLDATPEBEDCMENKGTGKAKATVYTPOEMAAQ
EPHSRLFTPEPNPRAAGLEKNVCRNPGDVGWCYTNPKLEFDICDVPQCAASSFD
CGKPOPEPKCGRVAGCVAYPRSWQWLSLRTGLMHPGGLTLPSPWVLTAAHL
EKSRSRSEFYKVLIGAHREYHLEPHVOIEVSKMSSEPARADILKLSPAIITDKVI
PCLSPSPNVADRTRECFITGMEGTGTYAGLKEARLPIVENKYCNVEFILTNTVK
TTELCAGHLAGTDSQDGSGLVCEPKDYIILGVTSMGLACARNPKRQYVAVSR
FVTHVEGVRRNN"
54..110
sig_peptide
/note="plasminogen signal peptide"
BASE COUNT 755 a 663 c 688 g 605 t
ORIGIN 2182 bp upstream of KpnI site.
Query Match 23.6%; Score 1178.8; DB 9; Length 2711;
Best Local Similarity 79.9%; Pred. No. 0;
Matches 1486; Conservative 0; Mismatches 292; Indels 82; Gaps 5;
QY 3159 TATGAGGCAATATCTTCACCACTGTACAGGAAGACCTGTCAATCTTGCTATCATG 3218
Db 900 TATGCTGGGAGATGTGGCTGTATCCGTGTGGCAGCACTGTGACGGCTGAGGCGCAG 959
QY 3219 ATACACACGTGCATCAGAGAGACCCAGAAACTACCCCAATGCTGCTGACCGAGAAC 3278
```

Db	2004	GAATATGAAGTATCTAATAGATGTTCTCGAGAGCCCCCAAGACACATATATGCTCTGCTAAAG	2063
QY	4359	CTAAGAGAGCCGCGCCGTATCATCAGACAAAGATATGACCTTGTCTGCGATCCCCAGAC	4418
Db	2064	CTAAGAGAGCTCTGCGATCATCTACACAAATATATCCACTTGTCTGCGATCCCCAAAT	2123
QY	4419	TACATGTCACCCGACGAGCATGTAATGTTATCATCTGCTGGGAGAAACCAGGTACC	4478
Db	2124	TATGTGCTGCTGACCGGACCGCAATGTTATCTGCTGGGAGAAACCAGGTACC	2183
QY	4479	TTTGGACAGTGCCTCTCAAGGAAGCCAGGCTCCTGTTATAGAAATGAATGTGCAT	4538
Db	2184	TATGGGGCTGGGCTTCTCAAGGAAGCCGCGCTCCGCTGATGAGAATGAAGTGTCAAT	2243
QY	4539	CACATATAGT-----ATATTGTGCTGAGCATTTTGGCC	4571
Db	2244	CGGTATGAGTCTTTCTGAATAGAAACATCAAAACACCGAGCTGTGCTGGACATTTGGCC	2303
QY	4572	AGAGGACATGACAGTTGCGCAGGCTGACAGTGGAGGGGCTTGTTGCTTCGAGAAGAC	4631
Db	2304	GGAGGACATGACAGTTGCGCAGGCTGACAGTGGAGGGGCTTGTTGCTTCGAGAAGAC	2363
QY	4632	AAATACATTTTCAAGGAGTCACTCTTGGGGTCTTGGCTGTGCACGCCCAATTAAGCT	4691
Db	2364	AAATACATTTTCAAGGAGTCACTCTTGGGGTCTTGGCTGTGCACGCCCAATTAAGCA	2423
QY	4692	GGTGCTATGCTGCTGCTTCAAGGTTTGTACTTGATTAAGGAATGATGAAATTAAT	4751
Db	2424	GGTGTATGCTGCTGCTTCAAGGTTTGTACTTGATTAAGGAATGATGAAATTAAT	2483
QY	4752	TAAATGGAGCGGA-GACAGAGTGAAGCATCACTACTTAATGAAGCTGAAGCTGGGTAAG	4810
Db	2484	TAAATGGAGCGGATTCACAGAGTGAAGCATCACTACTTAATGAAGCTGGGTAAG	2543
QY	4811	GATTTAGCATGCTGGAAATTAATAGACACATCAACAAAGAGACACTGCCAGCTACCA	4870
Db	2544	GATTTAGCATGCTGGAAATTAATAGACACATCAACAAAGAGAGTCCAGCTACCA	2603
QY	4871	GCTATGCCAAACCTTGGCATTTTGGTATTTTGTGTATAGCTTTAAGTCTGACAGA	4930
Db	2604	GGGAAACCAACCTTGGCATTTTGGTATTTTGTGTATAGCTTTAAGTCTGACAGA	2646
QY	4931	CAAAATCTGTATTAAGGTGATAGCTATGA-CAATTTGTTAAATAAATCACTGCACCTTA	4989
Db	2647	TGATTTCTGTATTAAGGTGATAGCTATGAACATTTGTTAAATAAATCACTGCACCTTA	2706
RESULT 9			
LOCUS	ARI05749	2296 bp	DNA linear PAT 14-FEB-2001
DEFINITION	Sequence 18 from patent US 6103244.		
ACCESSION	ARI05749		
VERSION	ARI05749.1	GI:12819814	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2296)		
AUTHORS	Donner, F., Schefflinger, F., Falkner, F., Gunter, and Pfeleiderer, M.		
TITLE	Methods for generating immune responses employing modified vaccinia		
JOURNAL	Of Fowlpox viruses		
FEATURES	Location/Qualifiers		
source	1..2296		
BASE COUNT	622 a 597 c 583 g 494 t		
ORIGIN	/organism="unknown"		
Query Match	22.2%;	Score 1110.6;	DB 6; Length 2296;
Best Local Similarity	81.4%;	Pred. No. 0;	
Matches 1340; Conservative	0; Mismatches 244; Indels 63; Gaps 2;		
Y	3159	TATCGAGGATATCTCCACCATCTGCACAGAGAGACCTGTCAATCTTGTCATCTATG	3218

Db 636 TATCGGGGAAATGTGGCTTTACCTTTCCGGGCAACCTCTGACCTGGATGACAG 695
Qy 3219 ATACCACTGGCATCAGAGACCCAGAAAATACCCAAATGCTGGCTCCAGAGAAC 3278
Db 696 ACCCTCAGACACATATACAGAGACAAAACCTTCCCTCAGAAAATTTGGATGAAAAC 755
Qy 3279 TACTGAGAAATCCAGATTTCTGGGAAACACCCCTGGTTTACACAGCCGATCCGTGTG 3338
Db 756 TACTGCGCAATCTGACGAGAAAAGGGCCCATGTGTCCATACACCAACAGCCAAAGT 815
Qy 3339 AGGTGGAGTACTGCAATCTGACACATGCTCAGAAAACGAATCGTGTCTAGAGACT 3398
Db 816 CGGTGGAGTACTGTAAGTATACCTGCTG----- 844
Qy 3399 CCCAGTGTGTTCCAGCTTCCAGCATGAGAGGCTCATTTGAAAGGACCCAGCTGAGCAA 3458
Db 845 -----TGACTCCTCCCAAGTATCCAGGACAAATTTGCTCCACAGCACTGAGCTA 899
Qy 3459 ACCCTGTGTGTCGGGAGTGTACCATGTAATGGCCAGTAATGAGCAATTTCTCC 3518
Db 900 ACCCTGTGTGTCAGAGAGTGTACCAAGTATGAGACAGCTACAGGCAATCTCTCC 959
Qy 3519 ACCAGTGTGACAGAGAGCATGTAATGTTGTCATGTCATGACACCAACCCGAGCAT 3578
Db 960 ACCACCAACCAAGAGAAAGTGTGATGTTGTCTATGACACCAACACCCGAGCAG 1019
Qy 3579 AGAGCCCCAGAAAATACCCAAATGATGCTGACAAATGAACTACTGAGAAATCCAGAT 3638
Db 1020 AAGACCCAGAAAATACCCAAATGCTGCTGACAAATGAACTACTGAGAAATCCAGAT 1079
Qy 3639 GCGGATACAGAGCCCTGTGTTTACATGAGACCCAGCATGAGTGGAGTACTGCAAC 3698
Db 1080 GCGGATACAGAGCCCTGTGTTTACATGAGACCCAGCATGAGTGGAGTACTGCAAC 1139
Qy 3699 CTGACGCGAGTGTGACACAGAGAGAGTGTGCTGCTCCGACTGTATCCAGGTT 3758
Db 1140 CTGAAAAAATGCTAGAGAAACAGAGAGTGTGAGACCTGCTGCTGCTGCTGCT 1199
Qy 3759 CCAAGCTAGAGGCTCTCTCTGACACAGACTATGTTGGAAATGGAAAGATACCG 3818
Db 1200 CCAAGTATGAGAGCTCTCTCTGACAGAGAGCTATGTTGGAAATGGAAAGATACCGA 1259
Qy 3819 GCGAAGAGGCAACCTGTTACTGAGAGCCATGCGAGAAATGGGCTCCAGAGAGCC 3878
Db 1260 GCGAAGAGGCGACACTGTACTGTGAGAGCCATGCGAGAACTGGGCTCCAGAGAGCC 1319
Qy 3879 CATGACACAGACGTTCAATCCAGGACAAATTAATGGAGGCTGAGAAAATTTAC 3938
Db 1320 CATGACACAGCAATTTTCACTCCAGAGCAATTCACGGGGGTGTGAAAAAATTTAC 1379
Qy 3939 TGGCGTAACCTGATGATGATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3998
Db 1380 TGGCGTAACCTGATGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439
Qy 3999 CTATTGACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4058
Db 1440 CTATTGACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
Qy 4059 CAAGTGAAGCCGAAAGAAATGCTGAGAGCAATTTAGGGGGGTGTGAGCCACCCCAT 4118
Db 1500 CAAGTGAAGCCGAAAGAAATGCTGAGAGGTTGTGGGGGGTGTGAGCCACCCCAT 1559
Qy 4119 TCGTGGCTGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4178
Db 1560 TCGTGGCTGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
Qy 4179 TTAATATATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4238
Db 1620 TTAATATATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
Qy 4239 TCAATCTCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4298

Db 1680 TCAATCTCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739
Qy 4299 GAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4358
Db 1740 GAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799
Qy 4359 CTAAGAGGCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4418
Db 1800 CTAAGAGGCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1859
Qy 4419 TACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4478
Db 1860 TACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1919
Qy 4479 TTTGGAGTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4538
Db 1920 TTTGGAGTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1979
Qy 4539 CACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4571
Db 1980 CACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2039
Qy 4572 AGAGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4631
Db 2040 AGAGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2099
Qy 4632 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4691
Db 2100 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2159
Qy 4692 GGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4751
Db 2160 GGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2219
Qy 4752 TAAATGAGCGGAGACAGAGTGAACA 4778
Db 2220 TAAATGAGCGGAGACAGAGTGAACA 2246

RESULT 10
114104 114104 2296 bp DNA Linear PAT 26-SEP-1995
LOCUS Sequence 18 from patent US 5445953.
DEFINITION 114104
ACCESSION 114104
VERSION 114104.1 GI:996527
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 2296)
AUTHORS
Dorner, F., Scheifflinger, F. and Falkner, F.G.
TITLE
Direct molecular cloning of a modified poxvirus genome
JOURNAL
Patent: US 5445953-A 18-29-AUG-1995;
FEATURES
Location/Qualifiers
1..2296
BASE COUNT 622 a 597 c 583 g 494 t
ORIGIN

Query Match 22.2%; Score 1110.6; DB 6; Length 2296;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1340; Conservative 0; Mismatches 244; Indels 63; Gaps 2;
Qy 3159 TATGAGGCAATATCTCCACACTGTCACAGAGAGAGCTGTCAATTTGGTCAATCATG 3218
Db 636 TATGAGGCAATATGCGGTGTATACGTTTCCGGGCAACCTGTCAACACTGGAGTGCACAG 695
Qy 3219 ATACCACTGGCATCAGAGAGCCAGAAAATACCCAAATGCTGGCTCCAGAGAAC 3278
Db 696 ACCCTCAGACACATATACAGAGACAAAACCTTCCCTCAGAAAATTTGGATGAAAAC 755
Qy 3279 TACTGAGAAATCCAGATTTCTGGGAAACACCCCTGGTTTACACAGCCGATCCGTGTG 3338
Db 756 TACTGCGCAATCTGACGAGAAAAGGGCCCATGTGTCCATACACCAACAGCCAAAGT 815
Qy 3339 AGGTGGAGTACTGCAATCTGACACATGCTCAGAAAACGAATCGTGTCTAGAGACT 3398
Db 816 CGGTGGAGTACTGTAAGTATACCTGCTG----- 844
Qy 3399 CCCAGTGTGTTCCAGCTTCCAGCATGAGAGGCTCATTTGAAAGGACCCAGCTGAGCAA 3458
Db 845 -----TGACTCCTCCCAAGTATCCAGGACAAATTTGCTCCACAGCACTGAGCTA 899
Qy 3459 ACCCTGTGTGTCGGGAGTGTACCATGTAATGGCCAGTAATGAGCAATTTCTCC 3518
Db 900 ACCCTGTGTGTCAGAGAGTGTACCAAGTATGAGACAGCTACAGGCAATCTCTCC 959
Qy 3519 ACCAGTGTGACAGAGAGCATGTAATGTTGTCATGTCATGACACCAACCCGAGCAT 3578
Db 960 ACCACCAACCAAGAGAAAGTGTGATGTTGTCTATGACACCAACACCCGAGCAG 1019
Qy 3579 AGAGCCCCAGAAAATACCCAAATGATGCTGACAAATGAACTACTGAGAAATCCAGAT 3638
Db 1020 AAGACCCAGAAAATACCCAAATGCTGCTGACAAATGAACTACTGAGAAATCCAGAT 1079
Qy 3639 GCGGATACAGAGCCCTGTGTTTACATGAGACCCAGCATGAGTGGAGTACTGCAAC 3698
Db 1080 GCGGATACAGAGCCCTGTGTTTACATGAGACCCAGCATGAGTGGAGTACTGCAAC 1139
Qy 3699 CTGACGCGAGTGTGACACAGAGAGAGTGTGCTGCTCCGACTGTATCCAGGTT 3758
Db 1140 CTGAAAAAATGCTAGAGAAACAGAGAGTGTGAGACCTGCTGCTGCTGCTGCTGCT 1199
Qy 3759 CCAAGCTAGAGGCTCTCTCTGACACAGACTATGTTGGAAATGGAAAGATACCG 3818
Db 1200 CCAAGTATGAGAGCTCTCTCTGACAGAGAGCTATGTTGGAAATGGAAAGATACCGA 1259
Qy 3819 GCGAAGAGGCAACCTGTTACTGAGAGCCATGCGAGAAATGGGCTCCAGAGAGCC 3878
Db 1260 GCGAAGAGGCGACACTGTACTGTGAGAGCCATGCGAGAACTGGGCTCCAGAGAGCC 1319
Qy 3879 CATGACACAGACGTTCAATCCAGGACAAATTAATGGAGGCTGAGAAAATTTAC 3938
Db 1320 CATGACACAGCAATTTTCACTCCAGAGCAATTCACGGGGGTGTGAAAAAATTTAC 1379
Qy 3939 TGGCGTAACCTGATGATGATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3998
Db 1380 TGGCGTAACCTGATGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439
Qy 3999 CTATTGACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4058
Db 1440 CTATTGACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
Qy 4059 CAAGTGAAGCCGAAAGAAATGCTGAGAGCAATTTAGGGGGGTGTGAGCCACCCCAT 4118
Db 1500 CAAGTGAAGCCGAAAGAAATGCTGAGAGGTTGTGGGGGGTGTGAGCCACCCCAT 1559
Qy 4119 TCGTGGCTGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4178
Db 1560 TCGTGGCTGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
Qy 4179 TTAATATATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4238
Db 1620 TTAATATATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
Qy 4239 TCAATCTCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4298

D	756	TATGCGCGAATCTCGACGGAAAAAAGGGCCCATGTGTGCATTACAAACCAACCCAGTGG	815
O	3339	AGGTGGGAGTACTGCATATCTGCACATCTGCTCAGAAACAGATATAGGTGTCTTAGAGACT	3339
D	816	CGGTGGAGTACTGTGTAAGATACCGTCTGT-----	844
O	3399	CCCACTGTGTTCACAGTTCACAGCATGAGAGGCTCATTTCTGAAGCAGCACCACTGGACAA	3458
D	845	-----TGAATCTCTCCCACTAGTATCCAGAGAAATATGGCTCCCAAGCAGCACCTAGCTA	899
O	3459	ACCCCTGTGTGTCGGCAGTGTCTACACATGGTAATGGCCAGAGTTATTCAGGACCATCTTCC	3518
D	900	ACCCCTGTGTGTCAGAGACTGTACACAGGTGATGACAGAGACTTCAGGSCATATCTTCC	959
O	3519	ACCACTGTACAGAAAGACATGTCTCAATCTTGTGTATCCATGACACACACCGGCTACAG	3578
D	960	ACCAACACACACAGAAAGAAATGTCACTTGTGTATCTATGACACACACCGGACACAG	1019
O	3579	AGGACCCCGAAATATACCCAAATGATGGCTGTACAAATGAATCACTACTSCAGAAATCCAGT	3638
D	1020	AAGACCCCGAAATATACCCAAATGATGGCTGTACAAATGAATCACTACTSCAGAAATCCAGT	1079
O	3639	GCGCATACAGGCCCTTGTGTATTACATGAGACACCCACATCAAGTGGAGTACTCCAC	3698
D	1080	GCGCATAAAGGCCCTGTGTATTACACAGACACCCACATCAAGTGGAGTACTCTCCAC	1139
O	3699	CTGACCCGATGTCTCAGACACAGAAAGGACTGTGTCTCTCTCCGACTGTCTATCCAGTT	3758
D	1140	CTGAAAAAATGTCTCAGAAACAGAAAGGAGTGTGTAAGACCTCTCCCTGTGTGTCTGT	1199
O	3759	CCAAGCCTAAGGCTCTCTCTTGAAACAAGACTGTATGTTGGAAATGGAAAGGATACCG	3818
D	1200	CCAAGTGTGAGACTCTCTCCGAAAGAACTGTATGTTGGAAATGGAAAGGATACCGA	1259
O	3819	GCGAATAAGGCAACACTGTTACTGCGGAGCGCCATCGACAGAAATGGGTGCCAGGAGCC	3878
D	1260	GCGAATAAGGCGACCACTGTACTGTGAGAGCCATCGACAGACTGGGTGCCAGGAGCC	1319
O	3879	CATAGACACAGACGTTATTCTCAGGACCAATTAATGGCAGGTCTGGAAAAAATTAC	3938
D	1320	CATAGACACAGATTTTACTCTCAGAGACCAATTCACGAGGGGTCTGGAAAAAATTAC	1379
O	3939	TGCGGTAAACCTGATGTGATCATATGTTGCTCTGCTACACAAATGAATCCAGAAAA	3998
D	1380	TGCGGTAAACCTGATGTGATGATGAGTGTGGTGTCTCTACACAAATTCAGAAAA	1439
O	3999	CTTTTGGATCTGTGATATCCCTCTGCTGATCCCTCTATTTGATTTGGGAAGCCT	4058
D	1440	CTTTAGACATCTGTGATGTCTCAGTGTGGGCCCTTATTTGATTTGGGAAGCCT	1499
O	4059	CAGGTGAGCCAGAGAAATCTCTCGAGACATTTAGAGGGGTGTGTGCCACCCACAT	4118
D	1500	CAGGTGAGCCAGAGAAATGTCTCGAGAGGGTGTGTGGGCCACCCACAT	1559
O	4119	TCTGTGCCCTGTGCAAGTCAAGTCTCAGACAAAGTTTGGAAAGCACTCTGTGGAGCAC	4178
D	1560	TCTGTGCCCTGTGCAAGTCAAGTCTCAGAAAGGTGTGTGGGCCACCCACAT	1619
O	4179	TTTAAATTTCCCAAGTGGGTGTGACATCGTCTGCACTGTGAAGTGTTCAGAGCT	4238
D	1620	TTTAAATTTCCCAAGTGGGTGTGACATCGTCTGCACTGTGAAGTGTTCAGAGCCT	1679
O	4239	TCAATCTTACAAAGTATCTTGGGTGACACACCAAGAGTGAACCTCGAATCTCATGTTTCAG	4298
D	1680	TCAATCTTACAAAGTATCTTGGGTGACACACCAAGAGTGAATCTCGAAGCCGATGTTTCAG	1739
O	4299	GAAATGGAAGTGTCTAGGCTTCTTGGAGGCCACACAGAAAGATATGCTTGCTTAAG	4358
D	1740	GAAATGGAAGTGTCTAGGCTTCTTGGAGGCCACACAGAAAGATATGCTTGCTTAAG	1799
O	4359	CTAAGCAGGCTGTGCGTGTACTGTACAAAGTATGCAAGTGTGTGTCATCCGAGAC	4418
D	1800	CTAAGCAGTCTGTGCGTGTACTGTACAAAGTATCCAGCTTGTGTCATCCCAAT	1859

OY	4419	TACATGCTCACGCCCGCCAGACTAATGTACATCACTGGCTGGGGAAGAACCCAAAGTACC	4478
Db	1860	TATGTGGTCCTGACCGGCACCGAATGTTTCACTGGCTGGGGAAGAACCCAAAGTACT	1919
OY	4479	TTTTGGGACTGGCCTTCTCAAGGAAAGCCCAAGCTCTTTTATTGAGAATAAGTGTCAT	4538
Db	1920	TTTGGACTGGCCTTCTCAAGGAAAGCCCAAGCTCTTGTATTGAGAAATTAAGTGTCAT	1979
OY	4539	CACATPAAAGT-----AFAFTTGTACTAGGATTTGGCC	4571
Db	1980	GCGTATAGTTTGTGAATGGAAGATCCAAATCCACCGAATCTGTGTCTGGCATTTGGCC	2039
OY	4572	AGAGGCACTGACAGTTTCCAGGGTGACAGTAGGAGGSCCTCTGTTTGTCTTCAAGAGAC	4631
Db	2040	GGAGGCACTGACAGTTTCCAGGGGTGACAGTAGGAGGTCTCTGTTTGTCTTCAAGAGAC	2099
OY	4632	AATATCAATTTTACAAGAGATCACTCTTGGGGGTCTTGGCTGTCACAGCCCCAATAAGCT	4691
Db	2100	AAATATCAATTTTACAAGAGATCACTCTTGGGGGTCTTGGCTGTCACAGCCCCAATAAGCT	2159
OY	4692	GGTGTATAGCTGTGCTTCAAGGTTTGTACTGTGATGAGGAATGATGAGAATAAT	4751
Db	2160	GGTGTATATTTGTGTTTCAAAGGTTTGTACTGTGATGAGGAATGATGAGAATAAT	2219
OY	4752	TAAATTGACGGAGACAGAGTAGMAACA	4778
Db	2220	TAAATTGACGGAGACAGAGTAGMAACA	2246
RESULT_11			
LOCUS	AX448883	2497 bp	DNA linear PAT 03-JUL-2002
DEFINITION	Sequence 1 from Patent WO0220813.		
ACCESSION	AX448883		
VERSION	AX448883.1	gi:21697781	
KEYWORDS	.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1		
AUTHORS	Cao Y.		
TITLE	Materials and methods relating to endothelial cell growth inhibitors		
JOURNAL	Patent: WO 0220813-A 1 14-MAR-2002;		
FEATURES	Karolinska Innovations AB (SE)		
source	Location/Qualifiers		
	1..2497		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	50..2482		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD37757.1"		
	/db_xref="gi:21697782"		
	/translation="MEHRERVLLLLFLKSGGEPLDVTVNTQASLFSYTKQLAG		
	STEECAKCEDDEPTCAROYHSKEQOCYIAMENRKSIIIRHADVLEEKRYLSE		
	CKTGNGKNRGTMKTNGITCOKMSSTPHRRPSPATHPSEGIEENYCANDPNDO		
	GMWCTTPDERKYDCDLIECEECHMGSGEYNDOKISKISGLCEAMQSOPAHAG		
	IYPSETPKKIKLKNYCNRPDELIPMCPTIDPNRKMYELCDIAPCTPPSPSSGPYCL		
	KDTGNIRGNVAITYVSGHTCOHMSADTPIITHRTDENPNCNDBENTCRNPDKRAPW		
	THGSKSQWSMTPRHROKTPENYPAUGLIMYCANPDADGPMCFITDSVREWCN		
	LAKCSGETASVAPPVVLLVDVEPSEEDCMFGNGGYRGKRATVYGTGPCDDMAAO		
	EBHRISITPETNPRAGILEKNYCRNPDGVGPMCVYTNNPKRLDYCDVQCAAPSFD		
	CKSPVERKCKGRVVGVCVAAPHSWPMOVSLRTFEGMHFCGGTLISPEWLTAAHL		
	ESKSPRSKYILGAHOENVLEPHVOELSVLSLPFRKRIALLXISPAVITLYDI		
	PACDSPHYVADBTREFCTIGMKETPGTGAGLLICADLPITENKYNCRKPYINGRO		
	STELCAAGGDCOGSDSGEPLVCFENDKTIILDGVISWLGCAARNPKFYIVYRSR		
	FVTVLEGVWRNN"		
BASE COUNT	692 a	624 c	636 g 545 t
ORIGIN			

Qy	3219	ATMCACTGGGACACAGAGACCCAGAAACTACCAAACTGCTGGCTGACCCAGAAC	3278
Dp	956	ACCCCTCACACATACAGAGACACCGAAACTCTCCCTGCAGAAATTTGGATGAAAC	10155
Qy	3279	TACTGCAGATCCAGATTCTGGGAAACAACCTGATGTACACACCGATCCGTGTG	33388
Dp	1016	TACTGCCGCAATCCAGAGGAAAAAGGCCCATGGTGCCTACACCAACGACCAAGTG	10757
Qy	3339	AGGTGGAGTACTGCAATCTACACATGCTCCAGAAACAGATCAGGTGTCTAGAGCT	33988
Dp	1076	CGGTGGAGTACTGTAAATACCGTCTCGT-----	11044
Qy	3399	CCCACTGTGTCCAGTCTCCAGATGAGAGGCTCAATCTGAGAGACCACTAGACAA	34585
Dp	1105	-----TGACTCCTCCCACTATCCAGGACAACTTGCTCCCAACGACCACTGAGCTA	11598
Qy	3459	ACCCCTGTGTCCGCACTGCTACCATGTAAATGCCCAGAGTATGAGGACACTTCTCC	3518
Dp	1160	ACCCCTGTGTCCAGAGATGCTACCATGTGTAGAGACAGACTCCAGGACACATCTCC	1219
Qy	3519	ACCACTGTACAGAAAGACATGCTCATTTGGTATCATGACACACACCGGACTCAG	3578
Dp	1220	ACCAACACACAGAAAGAAAGTGCATGTGTGATCATATGACACACACCGGACGAG	1279
Qy	3579	AGGACCCCGAAACATCCCAAAATGATGGCTGACAAATGAACTCTCAGAGATCCAGAT	36388
Dp	1280	AAAGCCCGAAAACTACCAAAATGCTGGCTGTACAAATGAATCTCAGAGAAATCCAGAT	1339
Qy	3639	GCCGATACAGGCCCTTGTTTATTCACATGAGCCCACTCAGATCAGTGGAGTACTGAC	3698
Dp	1340	GCCGATAAAGGCCCTTGTTTATTCACACAGACCCCACTCAGATGAGTACTGAC	1399
Qy	3699	CTACCCGATGCTCGACACAGAAAGGAGCTGTGTCGCTCCGCTGATCATCCAGGT	3758
Dp	1400	CTGAAAAATGCTCGAGAAAGAGAGATGTGTATGACACTCGGCTGTGTGCTGCTT	1459
Qy	3759	CCAAAGCTAGGCCCTCTCTGAAACAAGCTGTATGTTGGGATGGGAAAGATACCGG	3818
Dp	1460	CCAGATGTGAGACTCTCTCGAAAGAACTGTATGTTGGGATGGGAAAGATACCGA	1519
Qy	3819	GCGAAGACGCAACCACTGTACTTGGGAGAGCCATCCAGAAATGGCTGCCAGAGGCC	3878
Dp	1520	GCGAAGAGGCGCACTACTGTACTTGGGAGAGCCATCCAGAACTGGCTGCCAGAGGCC	1579
Qy	3879	CATAGACACAGACCTTATCCCGAGGACAAATTAATGGGAGGTCTGAGAAATTAATC	3938
Dp	1580	CATAGACACAGATTTTCTACTCCAGAGACAAATCCAGCGGGGCTCTGAGAAATTAATC	1639
Qy	3939	TGCCCTAACCCGTATGGTGACATCATGTCCCTGGTGTCTACACATGAATCCAGAAA	3998
Dp	1640	TGCCCTAACCCGTATGGTGAGTATGAGTGGTCCCTGGTGTCTACACAAATTAATTAATA	1699
Qy	3999	CTTTTGTACTGTGTATACCTCTCTGTGATCCTCTTATTTGATTTGTGGAGGCT	4058
Dp	1700	CTTTTGTACTGTGTATGTCTCTCTCTGATGTGGGCCCTCTTATTTGATTTGTGGAGGCT	1759
Qy	4059	CAGTGTGACCCCAAAATGTCCGGAAGCATTTGAGGGGGGTGTGGGCCACCCACT	4118
Dp	1760	CAGTGTGACCCCAAAATGTCTCGAAGAGTTTGAAGGGGGGTGTGGGCCACCCACT	1819
Qy	4119	TGCTGGCCCTGGCAAGTCACTCTAGAACAGGTTTGAAGGCACTTCTGTGGAGGACC	4178
Dp	1820	TGCTGGCCCTGGCAAGTCACTCTAGAACAGGTTTGAAGGCACTTCTGTGGAGGACC	1879
Qy	4179	TTAATTTCCCGAGATGGGTGCTGAGCTGTCTACTCTTTGAGAGAGTCTTAAAGCT	4238
Dp	1880	TTAATTTCCCGAGATGGGTGCTGAGCTGTCTACTCTTTGAGAGAGTCTTAAAGCT	1939
Qy	4239	TGATCTTACAAAGTCTCTGGGTGACACACAAAGTGAATGTGCAACCGCATTTTCAG	4298
Dp	1940	TGATCTTACAAAGTCTCTGGGTGACACACAAAGTGAATGTGCAACCGCATTTTCAG	1999
Qy	4299	GAAATGGAAGTCTAGGCTGTCTTGTGGAGCCCAACAAGCAGATTAATGCTTGTCAAG	4358

Db	2000	GAATTAAGATGCTCTAGGCTGTCTTCTTGGAGGCCACAGAAAAGATATATGCTCTGCTTAAG	2059
OY	4359	CTAAGAGGCTGCGCTCATCACTGACAAATGATGCCAGCTGTCTGCCATCCGAGAC	4418
Db	2060	CTAAGCAGTCTCTCCCTCATCACTGACAAATGATATCCACCTGTCTGCACTCCCAAT	2119
OY	4419	TACATGCTACCCGCGAGGACTGAATGTTATCATCACTGGCTGGGGAGAAAACCAAGGTACC	4478
Db	2120	TATGTGCTGCTGACCGGACCGAATGTTCTGTCACTGGCTGGGAGAAAACCAAGGTACT	2179
OY	4479	TTTGGGAGCTGGCCTTCTCAAGGAACCCAGAGTCTTGTPTTATGAATGAGTGGCAAT	4538
Db	2180	TTTGGAGCTGGCCTTCTCAAGGAACCCAGAGTCTTGTATATGAATGAATGATGTCMAAT	2239
OY	4539	CACATTAAGT-----ATATTGTGCTGACATTTGGCC	4571
Db	2240	CGCATATAGTTTCTGAATGGAAGTCCAAATCCACGAACTGTCTGTGGAGTTTGGCC	2299
OY	4572	AGAGGCACTACAGTTGGCCAGGGGTACAGTGGAGGGGCTCTGGTTTGTCTTCAGAAAGAC	4631
Db	2300	GGAGGCACTACAGTTGGCCAGGGGTACAGTGGAGGGTCTCTGTGTTGTCTTCAGAAAGAC	2359
OY	4632	AAATACATTTTAAAGAGATCACTCTTGGGGCTTGTGGCTGTGACAGGCCCAATAAGCT	4691
Db	2360	AAATACATTTTAAAGAGATCACTCTTGGGGCTTGTGGCTGTGACAGGCCCAATAAGCT	2419
OY	4692	GGTGTCTATGCTCTGTTTCAAGGTTTGTACTGTGATGAGGAATGATGAGAAATAT	4751
Db	2420	GGTGTCTATGCTCTGTTTCAAGGTTTGTACTGTGATGAGGAAGTGAATGAGAAATAT	2479
OY	4752	TAATTGGAGCGGAGACAG	4769
Db	2480	TAATTGGAGCGGAGACAG	2497
RESULT 13			
AR082437			
LOCUS	AR082437	2497 bp	DNA
LOCATION	Sequence 12 from patent US 5972896.		linear
ACCESSION	AR082437		
VERSION	AR082437.1	GI:10009163	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (phases 3 to 2497)		
AUTHORS	Davidson, D.J.		
TITLE	Antiangiogenic peptides and methods for inhibiting angiogenesis		
JOURNAL	Patent: US 5972896-A 12-26-Oct-1999.		
FEATURES	location/Qualifiers		
source	1..2497		
	/organism="unknown"		
BASE COUNT	691 a 624 c 637 g 545 t		
ORIGIN			
Query Match	22.1% Score 1104.8: DB 6; Length 2497:		
Best Local Similarity	81.4%: Pred. No. 0;		
Matches 1333; Conservative 0; Mismatches 242; Indels 63; Gaps	2;		
OY	3159	TATCGAGCATATCTCTCCACCACTGTCAAGGAAGACCTGTCAATCTTGGTCATCTATG	3218
Db	896	TATCGCGGGATGTGCTGTACCTGTCCGGGACACGTGACGACGTGGAGTGCACAG	955
OY	3219	ATACCACTGCGCATAGAGACCCAGAAAATACCAATCTGTGCTGACGAGAAC	3278
Db	956	ACCCCTGACACATATTAAGGACACACAAAATCTCCCTGCAAAAATTTGGATGMAAC	1015
OY	3279	TACTGCAAGAAATCAGATCTCTGGGAACAACCCGTGGTGTACACAAACGATCCGTGTG	3338
Db	1016	TACTGCGCGCATCTTACGAAAAAGGCCCCCATGTGTGCATCAACCAACGACCAAGTG	1075
OY	3339	AGGTGGAGTACTGCAATCTGACACAATGCTCAGAAAACAGAAATCAGGTGTCTTAGACT	3398

Db	1076	CGGTGGGAGTACTGTAAAGATACCGTCTCT	-----	1104
Oy	3399	CCCACTGTTTCCAGTTTCCAAACATGAGAGGCTATCTGTAGACAGACCACATGACGA		3458
Db	1105	-----TGACTCTCTCCCAAGTATCCACGAGACCAATTGGCTCCACACAGCACCACTGACCTA		1159
Oy	3459	ACCCCTGTGTCCGGCAGTCTACCATGTAAATGGCCAGAGTATCGAGGCACATTTTCC		3518
Db	1160	ACCCCTGTGTCTCAGAGACTCTACCATGTGTATGGACAGAGACTACGAGGCACATCTTCC		1219
Oy	3519	ACCACGTGCAGAGAAAGATGATGATTTGGTATCATGATGACACACACCGGATCAG		3578
Db	1220	ACCAACACACAGGAAGAAGTGTACGTGTTGGTATCTATGACACACACCGGCACAG		1279
Oy	3579	AGGACCCCAAGAAACTACCAATATGTGGCTTCACATGAACTATCTCAGAAATCCAAAT		3638
Db	1340	GCGCAATAAAGGCCCCCTGGTATTACACAGACCCACAGCTCAGTGTGGAGTACTGCAAC		1399
Oy	3699	CTCAGCCGATGTGTGACACAGAAAGGACTGTGTGGCTCTCCGATGTGCATCCAGGTT		3758
Db	1400	CTGAAAAAATGCTCAGAAACAGAAAGGAAATGTTGTAGACACTCCGCTGTGTCTGCTT		1459
Oy	3759	CCAAAGCTTAGGCTCTTCTGAAACAAGCTATATGTTGGGAATGGGAAAGATATCCGG		3818
Db	1460	CCAGATGTAGAACTCTCTCCGAAAGAACTGTATGTTTGGAAATGGGAAAGATATCCGA		1519
Oy	3819	GGCACAAAGCCAAACCATTGTACTGTGGAGCCCATCCAGAAATGGGTGCCCCAGAGAGCC		3878
Db	1520	GGCACAAAGGCGCACACACTGTACTGTGGAGCGCATCCAGAGACTGGCTGCCAGGAGAGCC		1579
Oy	3879	CATAGACACAGAGCTTCATCTCCAGGACAAATTAATGGAGGTCTGGAAAAAATTTC		3938
Db	1580	CATAGCACAGATTTTCTACTCCAGAGACAAATTCACAGCGGGGTCTGGAAAAAATTTC		1639
Oy	3939	TGCGCTAACCTGATGGTGAATCAATATGTTCCCTGGTCTACACATGAATCCAGAAA		3998
Db	1640	TGCGCTAACCTGATGGTGAATCAATATGTTCCCTGGTCTACACAAATCCAGAAA		1699
Oy	3999	CTTTTGACTACTGTGATATCCCTCTCTGTGCATCTCTTATTTGATTTGGGAAAGCT		4058
Db	1700	CTTTTGACTACTGTGATATCCCTCTCTGTGCATCTCTTATTTGATTTGGGAAAGCT		1759
Oy	4059	CAGTGTGAGCCGAAAGAAATGTCTCGAAGCATTTGAGGGGGGTGTGGGCCCAACCAT		4118
Db	1760	CAGTGTGAGCCGAAAGAAATGTCTCGAAGAGGTTGAGGGGGGTGTGGGCCCAACCAT		1819
Oy	4119	TCTGTGCTGTGCAAGTACTGTCAGAAACAGAGTTTGGAAAGCACTCTGTGTGAGGCAC		4178
Db	1820	TCTGTGCTGTGCAAGTACTGTCAGAAACAGAGTTTGGAAAGCACTCTGTGTGAGGCAC		1879
Oy	4179	TTAAATTTCCCAAGATGTGGTGTGACAGCGTCGCACTCTTGAAGAGTCTCTAAGGCT		4238
Db	1880	TTAAATTTCCCAAGATGTGGTGTGACAGCGTCGTCGCACTCTTGAAGAGTCTCTAAGGCT		1939
Oy	4239	TCAATCTACAAAGGTATCTCTGTGGTGCACACACAAGAGTGAACCTCGAATCTATGTTAC		4298
Db	1940	TCAATCTACAAAGGTATCTCTGTGGTGCACACACAAGAGTGAACCTCGAATCTATGTTAC		1999
Oy	4299	GAAATAGAGTGTCTAGGCTGTCTCTGTGGAGCCCAACAAGACAGATATTTGGCTTGTCTAAG		4358
Db	2000	GAAATAGAGTGTCTAGGCTGTCTCTGTGGAGCCCAACAAGATATTTGGCTTGTCTAAG		2059
Oy	4359	CTAAGAGGCTCGCGCTATCTACAGCAAAAGTAAATGCAAGTGTGTCTGCCATCTCCAGAC		4418
Db	2060	CTAAGAGTCTCTCGCGTATCTACAGCAAAAGTAAATGCAAGTGTGTCTGCCATCTCCAGAT		2119
Oy	4419	TACATGTACCGCCAGGACTGAATGTTACATCACTGGCTGGGAGAAACCAAGGTAC		4478

QY	2120	TAATGTGGTCGTGACGGAGCCGAATGTTTCGTCACTGGCTGGGAGAAACCAAGTACT	2179
QY	4479	TTTGGGACTGGCCTTCTCAAGAGAACCCACGCTCTTGTATTAGANTGAAGTGTCAAT	4538
Db	2180	TTTGGACCTGGCCTTCTCAAGAGAACCCACGCTCTTGTATTAGANTGAAGTGTCAAT	2239
QY	4539	CACATAAGT-----ATATTGTGCTGAGCAATTTGGCC	4571
Db	2240	CGCATATGTTTTCGATGGAAAGATGCCAATCCACGCAACTCTGTGCTGGGCAATTTGGCC	2299
QY	4572	AGAGGCACTGACAGTTGGCCAGGGTGCACGTGAGAGGCGCTCTGTTGCTTCGAGAGGAC	4631
Db	2300	GGAGGCACTGACAGTTGGCCAGGGTGCACGTGAGAGTCTCTTGTTGCTTCGAGAGGAC	2359
QY	4632	AAATPACATTTTACAGAGAGTCACTTCTTGGGCTCTTGCTGTGACGCCCAATTAAGCT	4691
Db	2360	AAATPACATTTTACAGAGAGTCACTTCTTGGGCTCTTGCTGTGACGCCCAATTAAGCT	2419
QY	4692	GGTGTATGCTGTGTTTCAAGTTTGTACTTGGATTGAGGAAATGAGAAATTAAT	4751
Db	2420	GGTGTATGTTGTGTTTCAAGTTTGTACTTGGATTGAGGAGATGATGAGAAATTAAT	2479
QY	4752	TAATTGACGGGAGACAG 4769	
Db	2480	TAATTGACGGGAGACAG 2497	
RESULT 14			
LOCUS	AR085163	2497 bp	DNA
DEFINITION	Sequence 12 from patent US 5981484.		
ACCESSION	AR085163		
VERSION	AR085163.1	GI:10011933	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 2497)		
TITLE	Davidson, D.J.		
JOURNAL	Antigenic peptides and methods for inhibiting angiogenesis		
FEATURES	Patent: US 5981484-A 12 09-NOV-1999;		
source	location/Qualifiers		
	1..2497		
	/organism="Unknown"		
BASE COUNT	691 a 624 c 637 g 545 t		
ORIGIN			
Query Match	22.1%;	Score 1104.8;	DB 6; Length 2497;
Best Local Similarity	81.4%;	Pred. No. 0;	
Matches 1333; Conservative	0; Mismatches 242;	Indels 63;	Gaps 2.
QY	3159	TATGAGGCAATCTCTCACCACTGTACAGAGAGAGGACCTGTCAATCTTGTATATG	3218
Db	896	TATGCGGGGAATGTGGCTTACCGTGTCCGGGACACCTGTGACACTGAGTGCACAG	955
QY	3219	ATACCACACTGGGATTCAGAGAGCCGACGAAACCTGCCAAATGCGTGGGCTGTACCGAGAAC	3278
Db	956	ACCCTTCACACATTAACAGAGAACCAAGAAATCTCCCTGCAAAATTTTGATGAGAAAC	1015
QY	3279	TACTGCAGGAATCCAGATTCTGGGAAACAACCCGTGGTTTACACACCGATCGGTGTG	3338
Db	1016	TACTGCCCCCATCTGTGACGAGAAAGGCCCCCATGTGTGCATATACACCAACAGCCAGCA	1075
QY	3339	AGTGGGAGTACTCAATCTGACACATATCTCAGAAACAGATCAGGTGTCTAGAGACT	3398
Db	1076	CGGTGGAGTACTGTAGATACCGTCTG-----	1104
QY	3399	CCCACTGTTTTCAGTTCCACAGATGAGGCGCTCAATTCGAAAGACCAACACTGAGCA	3458
Db	1105	-----TGACTCTTCCCAAGTATCAGAGACAATTGGCTCCACACGACCACTGAGACTA	1159
QY	3459	ACCCCTGTGTGCTGCGGAGTGTACATGGTTAATGGCAGAGTTATCGAGGCATTTCTC	3518

Db	1160	ACCCTGTGTGTCAGAGACTGCTACACGATGGATGGACAGACTACGAGGACATCTCTCC	1219
Qy	3519	ACCACTGTACAGAGAAAGACATCTGTCAATCTGTGTATCATATGACACACACCGGGATCTAG	3578
Db	1220	ACCAACACACACAGGAAAGAAAGATGTCACTCTGTGGTATCTATGTACACACACCGGGACGAG	1279
Qy	3579	AGGACCCCGAAAACATACCACAAATGATGGCTGTACAAATGAACATCTACTGCAGAAATCCAGAT	3638
Db	1280	AAGATCCCGAGAAACATACCACAAATGGCTGTACAAATGAACATCTACTGCAGAAATCCAGAT	1339
Qy	3639	GCGCATACAGGCGCTTGGTGTATTACATGAGACCCACATCAGATGGAGTACTGTCAAC	3698
Db	1340	GCGGATTAAGGCGCCCTGTGTATTACACAGACAGACCCACATCAGATGGAGTACTGTCAAC	1399
Qy	3699	CTTAGCCGATGCTCAGACACAGAAAGGACTGTGTCTCTCTCCGACTGTATCAGAGTT	3758
Db	1400	CTTAAAAAATGTCTCGAAGAACAGAAAGCAGTGTATGAGACATCTCCGCTGTGTCTCGGTT	1459
Qy	3759	CCAAGCCTAGGGCCCTCTTTGTAACAAAGACTGTATGTTTGGGAATGGGAAGGATACCGG	3818
Db	1460	CCAGATGTAGAGACTCTTCCGAAAGAACTGATGTTTGGGAATGGGAAGGATACCGA	1519
Qy	3819	GGCAAAAGGCAACACAGTTCAGTGGAGAGCCCATCCAGAGAAATGGGTGTGCCAGAGGCC	3878
Db	1520	GGCAAAAGGCGACCACTGTTACTGGACCCATGTCCAGACTGGGTGTGCCAGAGGCC	1579
Qy	3879	CATAGACACAGCAGCTTCATTCAGAGGACAAATTAATGGCAGGCTGTGAAAAAATTTAC	3938
Db	1580	CATAGACACAGCATTTTCTACTCCAGAGACAAATTCACAGCGGGCTGTGAAAAAATTTAC	1639
Qy	3939	TGCGCTTAACCTGATGGAGACATCAATAGTGCCTGTCTCTACAAATGAATCCAAAGAAA	3998
Db	1640	TGCGCTTAACCTGATGGAGATGATAGAGGGGCTGTGTCTCTACAAATGAATCCAAAGAAA	1699
Qy	3999	CTTTTGTACTACTGTGAATACCCCTCTGTGTGATCCTCTCATTTGTATGTGTGGGAAGCT	4058
Db	1700	CTTTTGTACTACTGTGAATCTCTCTGTGTGATCCTCTCATTTGTATGTGTGGGAAGCT	1758
Qy	4059	CAAGTGAGCCGAGAAAATGTCTGTGAAGCATTTGAGGGGGTGTGTGCCCCCACCACAT	4118
Db	1760	CAAGTGAGCCCGAAMAAATGTCTGTGAAGGGTTGAGGGGGTGTGTGCCCCCACCACAT	1819
Qy	4119	TCCGTGCCCCGTGCAGTCACTCAAGTCAAGCAAGAGTTTGAAGACACTCTGTGGAGGACAC	4178
Db	1820	TCCGTGCCCCGTGCAGTCACTCAAGTCAAGAGTTTGAAGATGCACTCTGTGGAGGACAC	1879
Qy	4179	TTTAAATCCCCAGATGGGTGTGCATCTGCTGTCACTGCTGTGAAGAGTCTTAAGGCT	4238
Db	1880	TTTAAATCCCCAGATGGGTGTGTATCTGTCTGTGAGCCACTCTGTGGAGAGTCCCAAGGCT	1939
Qy	4239	TCATCTTACAGGTATCTCTGGGTGCACACCAAGAGTGAACTCGAATCTCATGTTTCAG	4298
Db	1940	TCATCTTACAGGTATCTCTGGGTGCACACCAAGAGTGAAATCTCGAAGCGCATGTTTCAG	1999
Qy	4299	GAAATAGAGTGTCTAGGCTGTCTTGGAGCCCAACAAAGATATGTGCTTGCTGCTTAAG	4358
Db	2000	GAAATAGAGTGTCTAGGCTGTCTTGGAGCCCAACAAAGATATGTGCTTGCTGCTTAAG	2059
Qy	4359	CTAAGCAGGCGTGGCTGTATCATCTGACAAAGTAATGCCAGTGTGTCTGCATCTCCAGAC	4418
Db	2060	CTAAGCAGTCTGGCTGTATCTATCTACGAAAGTAATCCAGGTTGTCTGCATCTCCCAAT	2119
Qy	4419	TACATGTACCGCCAGAGACTGATGTTTACATCACTGGCTGGGAGAAACCCAAAGTATAC	4478
Db	2120	TATGTGTGTCTGACCGGACCGAATGTTTCTCATCTGGCTGGGAGAAACCCAAAGTACT	2179
Qy	4479	TTTGGGACAGGCGCTTCTCAAGAGAAAGCCAGGCTCTGTATTTAGTAAGTAAGGTGCAAT	4538
Db	2180	TTTGGAGCTGTGGCTTCTCAAGAGAAAGCCAGGCTCTGTGTGATTTAGATTAAGTGTGCAT	2239
Qy	4539	CACATATAGT-----ATATTTGTGCTGACATTTGGCC	4571
Db	2240	CGCTTAAGATTTTCTGAATGAAGATCTCAATCCACGAAATCTGTCTGTGGGCAATTTGGCC	2299

QY	4572	AGGGCACTGACAGTGGTCCAGGGGTGACATGGAGGGGCTCTGGTTGCTTGAGAAAGAC	4631
Db	2300	GGAGGCACTGACAGTGGTCCAGGGGTGACATGGAGGGTCTCTGGTTGCTTGAGAAAGAC	2359
QY	4632	AAATACATTTTACAGAGATGACCTTTTGGGCTCTGGCTGTGCACGCCCAATTAACCT	4691
Db	2360	AAATACATTTTACAGAGAGTCACTTTCTGGGCTCTGGCTGTGCACGCCCAATTAACCT	2419
QY	4652	GGTGTCTATGCTGTGTGTTTCAGGTTTGTACTTGGATTGAGGAAATGATGAGAAATAT	4751
Db	2420	GGTGTCTATGCTGTGTGTTTCAGGTTTGTACTTGGATTGAGGAGATGATGAGAAATAT	2479
QY	4752	TAATTGGACGGGAGACAG	4769
Db	2480	TAATTGGACGGGAGACAG	2497

RESULT 15				
HOMPLASM				
LOCUS	HOMPLASM	2497 bp	mrna	linear
DEFINITION	Human plasminogen mRNA, complete cds.			
ACCESSION	M74220			
VERSION	M74220.1	GI:190025		
KEYWORDS	plasminogen; zymogen;			
SOURCE	Homo sapiens female adult liver cDNA	to mRNA.		
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Browne,M.J., Chapman,C.G., Dodd,I., Carey,J.E., Lawrence,G.M.P., Mitchell,I.D. and Robinson,J.H.			
TITLE	Expression of recombinant human plasminogen and aglycoplasminogen in HeLa cells			
JOURNAL	Fibrinolysis (1991) In press			
FEATURES	Location/Qualifiers			
source	1..2497			

CDS

```

      1  /db_xref="taxon:9606"
      2  /haplotype="diploid"
      3  /sex="female"
      4  /tissue_type="liver"
      5  /dev_stage="adult"
      6  50..2482
      7  /codon_start=1
      8  /product="plasminogen"
      9  /protein_id="AA36451.1"
     10  /db_xref="GI:190026"
     11  /translation="MEHKEVYLLLLFLKSGCEPLDDYVNTQGSLSFSYTKKQLAGA
     12  SLIEBCKACEDEEFTCAFOYHSKEQOCVIMENKRSIIIRMDVLFKKYVLST
     13  CKTGKNCNRTGMSKTKNGITICQKWSSTSPHRRFSPAHPSGELEENYCRNPDPC
     14  GWCYTTPDEKRYDYCDILBCEECMCSEENIDGISTMSGLECAMDSQSPHAGH
     15  YLSPFPNKLKNCNCRNPDLPKCFITDPKRNELDIDICTRTPPSSGPTQCC
     16  KGTGEYRGCVAVTVSGTHCOHMSAQPHTHNTPENFPCKNIDENYCNPKGKAPC
     17  KGTNSOYRWEMCKIPSCDSSVSTEQALPTAPPELPIVQDCHDGGQSYGSGTST
     18  TKGKCOGSSSWTPEHRHCKPENYPMNAIGMNCNPDMDKGMCFETDPSRYMGE
     19  LKCKSGTESVAVAPPPVLLIDPVETPESEECMGNGKYGKRAITVYGTGTPQMAA
     20  EHRHSITPEETNPAGLEKKNYCNPDGCGWCYTKNIPDKLYIYCDYCOAASPT
     21  CGNFOPEKCKPGRVYGCVAHPHSMQVSLTRGMMFCGTLISPEWVITDKIC
     22  EKSPPRSKTYLLGAHQENLEPHVOELVSKLFLPETRDLALKLSAPVATDKAC
     23  PACLBSPTNVVADRTCECVTGMGTGTGTFAGILKRAQLPIENYCNKTEFLNRSV
     24  SIELAGHLAGDTGSCDQSDSGPLVCFEKDKYILDGVTSMLGCAQPNKPGCVYVSS
     25  FVTYIEGVWRNN"
     26  50..106
     27  /product="plasminogen"
     28  107..2479
     29  /product="plasminogen"
     30  624 c 637 g 545 t

```

Query Match 22.1%; Score 1104.8; DB 9; Length 2497;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 242; Indels 63; Gaps 2.

```

QY 3159 TATGAGGATATCTCTCCACACCTGTGCAGAGAGACCTGTGATCTTGTGATCTATG 3218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 896 TATCGCGGGAATGTGTGCTTACCGTGTCCGGACACACTGTGATGACCTGCACAG 955
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3219 ATACACACCTGTCATGAGAGAGACCCGAGAAAATACCCAAATGCTGGCTGCAGAGAAC 3278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 956 ACCCTGCACACATPAACAGAGACACGAGAAAATCTTCCCTCGCAAAAATTTTGGATGAAAAC 1015
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3279 TATCGAGAGATCCAGATCTTGGGAAACAACTGGGTGTACCAACCATCCGTGTG 3338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1016 TACTCGCCGATCTGCAGGAGAAAAGGGCCCTGGGTCCATCAACCAACAGCCAAAG 1075
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3339 AGGTGGAGTACTGCAATCTGACAAATGCTCAGAAAACAGAAATCAGGTGTCTAGAGACT 3398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1076 CGGTGGAGTACTGCAATGATACCTGCTG----- 1104
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3399 CCCACGTGTGTTCCAGTTCCAGAGTGAAGTCAATTTGAGACAGCACCAATGAGCAA 3458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1105 -----TGACTCTCCCGCCAGTATCCACGGAACAAATGGCTCCACACACCTGAGCTA 1159
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3459 ACCCTGTGTGTCGGGAGTGTCTACCATGTATGATGCCAGATTATGAGGACATTTCTC 3518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1160 ACCCTGTGTGTCAGAGTGTCTACCATGTATGATGACAGAGCTACCGAGGACATCTCTC 1219
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3519 ACCACTGTACAGAGAGACATGCAATCTGTGTCATGTCATGACACACAGCGGACATG 3578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1220 ACCACACACAGAGAAAGTGTCAATCTTGTGTCATGTCATGACACACAGCGGACACG 1279
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3579 AGGACCCCAAGAAAATCTACCAATGATGCTGTGACAAATGAACTACTGTGAGAAATCCAGAT 3638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1280 AAGACCCCAAGAAAATCTACCAATGATGCTGTGACAAATGAACTACTGTGAGAAATCCAGAT 1339
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3639 GCCGATACAGGCGCTTGTGTTTACCATGAGACCCAGATCAGAGTGGAGATGACTGCAAC 3698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1340 GCCGATACAGGCGCTTGTGTTTACCATGAGACCCAGATCAGAGTGGAGATGACTGCAAC 1399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3699 CTGAGCGGATGTCAGAGACAGAGAGAGTGTGCTCTCTCCAGTGTCAATCCAGTT 3758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1400 CTGAAAAAATGCTCAGAGAGAGAGAGAGTGTGTACACCTCCGCTGTCTGCTGCTT 1459
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3759 CCAAGCTTAGGCGCTCTCTTGAAACAAGACTGTATGTTGGAGATGGAAAGATACCG 3818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1460 CCAAGATGTAGAGACTCTCTCGAAGAAAGATGTATGTTGGAAATGGAAAGATACCGA 1519
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3819 GCGAAGAGGACACACACTGTACTGTGAGAGCCATGCCAGAAATGGCTGCCAGAGAGCC 3878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1520 GCGAAGAGGCGACACACTGTACTGTGAGAGCCATGCCAGAAATGGCTGCCAGAGAGCC 1579
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3879 CATAGACAGACAGTCTCATTCAGAGGACAATGAATGAGGACAGTCTGAGAAAAAATATAC 3938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1580 CATAGACAGACAGTCTCATTCAGAGGACAATGAATGAGGACAGTCTGAGAAAAAATATAC 1639
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3939 TGCCTTAACCTGTGATGTGATCAATGTGCTGCTGTGCTGTACACATGAATCCAGAAAA 3998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1640 TGCCTTAACCTGTGATGTGATGTGAGTGTGCTGCTGTGCTGTACACATGAATCCAGAAAA 1699
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3999 CTTTTGACTACTGTGATATCCCTCTCTGTGTCATCTTTCATTTGATTGTGGAGAGCT 4058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1700 CTTTGAGTACTGTGATGTGATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1759
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4059 CAAGTGTGAGCCGAGAAATGTCTGTGAGAGCAATTTAGSGGGGTGTGGCCACCAAT 4118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1760 CAAGTGTGAGCCGAGAAATGTCTGTGAGAGCAATTTAGSGGGGTGTGGCCACCAAT 1819
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4119 TCTGTGCTGTGCAAGTCTCTCAGAAACAAGTTTGGAAAGCACTTGTGAGAGCACC 4178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1820 TCTGTGCTGTGCAAGTCTCTCAGAAACAAGTTTGGAAAGCACTTGTGAGAGCACC 1879
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4179 TTAATATCCCAAGATGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1880 TTAATATCCCAAGATGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1939
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4239 TCAATCTAAGATCATCTGTGGTGTGACACCAAGAAATGAACCTGCAATCTCATGTTGAG 4298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 1940 TCAATCTAAGATCATCTGTGGTGTGACACCAAGAAATGAATCTCAACCGCATGTGAG 1999
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4299 GAAATTAAGATGTCTAGAGCTGTCTGTGAGCCACACAGACAGATATTCCTTTGCTAAG 4358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2000 GAAATTAAGATGTCTAGAGCTGTCTGTGAGCCACACAGAAATATTCCTTTGCTAAG 2059
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4359 CTAAGAGGCGCTGTGCTGATGACAGCAAGTATGTCACCTGTGTCGCATCCAGAC 4418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2060 CTAAGAGTCTGTGCTGATGACAGCAAGTATGTCACCTGTGTCGCATCCAGAC 2119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4419 TACATGTACCGCGAGACAGTGAATGTATCACTGCTGTGAGAGAAACCAAGTAC 4478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2120 TATGTGTGTGTGACCGGACCGAATGTTCCTCACTGTGTGTGAGAGAAACCAAGTACT 2179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4479 TTTGGAGTGTGCTTCTCAAGAGACCGAGCTGTGTATGAGATGAATGATGAT 4538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2180 TTTGGAGTGTGCTTCTCAAGAGACCGAGCTGTGTATGAGATGAATGAATGATGAT 2239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4539 CACTATAGT-----ATATTGTGCTGAGCATTTGGCC 4571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2240 CGCTATGATGTTTCTGAATGAAAGTCCATTCACCGAAGCTGTGTGTGAGCATTTGGCC 2299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4572 AGAGGACTGACAGTTTGCAGAGGTGACAGTGTGAGAGGCTGTGTGCTTGTGCTGAGAGAGAC 4631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2300 GAGGACTGACAGTTTGCAGAGGTGACAGTGTGAGAGGCTGTGTGCTTGTGCTGAGAGAGAC 2359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4632 AATATATTTTCAAGAGAGTCACTCTGTGGGTGTGTGCTGTGAGAGGAAATGATGAGAAATAT 4691
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2360 AATATATTTTCAAGAGAGTCACTCTGTGGGTGTGTGCTGTGAGAGGAAATGATGAGAAATAT 2419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4692 GGTGTATGCTGTGTCAAGGTTTGTACTTGAATGAGGAAATGATGAGAAATAT 4751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2420 GGTGTATGCTGTGTCAAGGTTTGTACTTGAATGAGGAAATGATGAGAAATAT 2479
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4752 TAAATGACGGAGACAG 4769
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2480 TAAATGACGGAGACAG 2497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: March 6, 2003, 12:17:38
 Job time : 12418 secs

FT	/tag- f
FT	/number=1
FT	/note="Klingler"
FT	2893..3128
FT	/tag- g
FT	/number=2
FT	/note="Klingler"
FT	3162..3398
FT	/tag- h
FT	/number=3
FT	/note="Klingler"
FT	3468..3704
FT	/tag- i
FT	/number=4
FT	/note="Klingler"
FT	3774..4032
FT	/tag- j
FT	/number=5
FT	/note="Klingler"
FT	4022..4024
FT	/tag- k
FT	/note="start" PMGN protein protease"
FT	4074..4106
FT	/tag- l
FT	/note="PMGN protein protease activation ARG"
FT	4769..4997
FT	/tag- m
FT	4997..6020
FT	/tag- n
XX	
PN	W09013640-A.
XX	
PD	15-NOV-1990.
XX	
PE	26-APR-1990; 90MO-US02296.
XX	
PR	01-MAY-1989; 89US-0345801.
XX	
PA	(UYNO-) UNIV NOTRE DAME DU.
XX	
PL	Caetellino FJ, Whitefleet-Smith JL, Rosen ED, McInden JH;
DR	WPj: 1990-361475/48.
DR	P-PSDB: AAR08065.
XX	
PT	Expression of human plasminogen in eukaryotic cell vector - used
PT	in thrombolytic therapy
XX	
PS	Disclosure; Fig 2; 110pp; English.
XX	
CC	Eukaryotic cell expression vectors, pref. insect cell expression
CC	vectors, esp. a baculovirus vector, chosen from the gp. of
CC	Autographa californica and Bombyx mori nuclear polyhedrosis virus
CC	reactors, or SV40-, polyoma-, adeno-, VSV or BPV virus, contg. this
CC	gene, can be introduced into hosts, eg. a Spodoptera frugiperda cell.
XX	
SQ	Sequence 6020 BP; 1612 A; 1502 C; 1482 G; 1424 T; 0 other;
Query Match	24.8%; Score 1238.4; DB 11; Length 6020;
Best Local Similarity	81.1%; Pred. No. 0;
Matches 1511; Conservative	0; Mismatches 271; Indels 80; Gaps
OY	3159 TATGAGGCATATCTTCACCACTGTGCACGAGAAGACTGTCAATCTTGTCATCTATG 3218
Db	3186 TATGGCGGGAATGTGCTGTATTACCGTGTCCGGCACACCTGTACACACTGCGAGTCACAG 3245
OY	3219 ATACCACACTGCATCAGAGACCCCAGAAAATACCAAATGCTGACCGAGAAC 3278
Db	3246 ACCCTGCACACATTAACAGACACCAAGAAATCTCCCTCGAAAAAATTTGGATCAAAAC 3305
OY	3279 TACTGCAGGAATCCAGATTGTGGGAACAACCTGSGTGTTCACACACCGATCCGTGTG 3338
Db	3306 TACTGCCCATCTCTGACGGAAGAGCCCCCATGTGTCATACCAACCAAGCCAACTG 3365

QY	3339	AGTGGAGTACTGCAATCTGACACAATGCTCAGAAAACGAATCAAGTGTCTTAGACACT	3338
Db	3366	 CGGTGGAGGTACTGTAAATACGCTCTG-----	3394
QY	3399	CCCACTGTTGTCAGATTCACAGCTAGAGGGCTCATTTGAAAGCAGCACCAACTGAGCA	3458
Db	3395	 -----TAGCTCTCTCCCAAGTATCCACGSAACAATTGGCTCCACAGACCACTGAGCTA	3449
QY	3439	ACCCCTGTGGTCGGGACAGTGTCTACCATGGTAATGGCCAGAGTTATTCAGAGGCATTTCC	3518
Db	3450	ACCCCTGTGGTCGACAGGACTCTACCATGGTGAATGGACACAGATACCGAGGGACATCTCC	3509
QY	3519	ACCACTGTCAAGAGNAGAGACATGTCTCATCTGGTCATCCATGACACCAACCCGGCATCG	3578
Db	3510	 ACCAACACACAGGAAGAAGATGTCAGTCTGGTCACTATGACACCAACCCGGCACCG	3569
QY	3579	AGGACCCCAAAAACCTACCAATGATGGCTGTACATAATACTACTGCAAGAAATCCAGAT	3638
Db	3570	AAGACCCCAAAAACCTACCAAAATGGCTGGCTGACATATACTACTGCAAGAAATCCAGAT	3629
QY	3639	GCCCATACAGGCCCCGTGTGTTTACCATGAGACCCAGCATAGGTGGAGAGTACGCAAC	3698
Db	3630	GCCGATTAAGGCCCCCTGTGTGTTTACCAACAGACCCAGGCTCAGGTGGAGAGTACGCAAC	3689
QY	3699	CTGACGCGATCTCTAGACACAGAAAGGAGTCTGTGCTCTCCGATCTCATCCAGTT	3758
Db	3690	CTGAAAAATCTCTAGGAACAAGAAAGAGTGTGTAGCACTCCCTCTGTGTCTGCTT	3749
QY	3759	CCAAAGCCTAGGGCCCTCTTGTGAACAAGACTATGTTTGGAAATGGAAAGATACGG	3818
Db	3750	CCAGATGTAGAGACTCCTCTCCGAAGAAAGACTATGTTTGGAAATGGAAAGATACGGA	3809
QY	3819	GGCAAGAAAGCAACACTGTTTACCTGGGAGACCCATGSCAGAAATGGCTCCAGAGGCC	3878
Db	3810	GGCAAGAGGGGAGCACTGTTACTGGGAGCCCATGCGCAGACTGGGGCTGCCAGGAGGCC	3869
QY	3879	CATGACACACAGCAGTTCAATCCAGGAGACAATAATATGGGCAAGTCTGGAAAAAATTAC	3928
Db	3870	CATGACACACACTTATTTACTCTCCAGAGACAAATCCACGGGCGGGTCTGGAAAAAATTAC	3929
QY	3939	TGCCGTAAACCTGTATGGGTGACATCAATGAGCCCTGGTGTACACATCAATCCAAAGAA	3998
Db	3930	TGCCGTAAACCTGTATGGGTGATGTAGTGGTGGTCCCTGGTGTACACGACAATCCAAAGAA	3989
QY	3999	CTTTTGTACTCTGTATTCCTCTCTCTGTGCACTCTTTCAATTTGATTTGTSGAAAGCT	4058
Db	3990	CTTTACGACTACTGTGATGTCCCTCAGTGTGCGGCCCTTCAATTTGATTTGTSGAAAGCT	4049
QY	4059	CAAGTGGAGCCGAGAAATCTCTCGAAGACTGTAGGGGGGTGTGTGGCCACCACCACAT	4118
Db	4050	CAAGTGGAGCCGAGAAATGTCCTCGAAGGGTGTAGGGGGGTGTGTGGCCACCACCACAT	4109
QY	4119	TCCTGGCCCTGGCAAGTCACTCTCAGACAAAGTTTGGAAAGCACTTCTGTSGAGGACCC	4178
Db	4110	TCCTGGCCCTGGCAAGTCACTCTTAGAACAAAGTTTGGAAATCACTTCTGTSGAGGACCC	4169
QY	4179	TTAATATCCCCAGAGTGGGTGTGACAGTCTGCTCACTGCTTAAAGAGCTCCAGAGCCT	4238
Db	4170	TTGATATCCCAAGATGGGTGTGACAGTCTGCTGCCACATGCTTSGAAGAGTCCCAAGGCT	4229
QY	4239	TCATCTTACAAGTCACTCTGGGTGCAACACCAAGAAAGTAACTCTGAAATCTCATGTTAC	4298
Db	4230	TCATCTTACAAGTCACTCTGGGTGCAACACCAAGAAAGTAACTCTGAAACCGCATGTTAC	4289
QY	4299	GAATATAGAAAGTGTAGAGGCTTCTTGGAGCCCAACAGCAGAGATATTTGGCTGTAAAG	4358
Db	4290	GAATATAGAAAGTGTAGAGGCTTCTTGGAGCCCAACAGAAAGATATTTGCTGTGTAAAG	4349
QY	4359	CTAAGCAGGCTCGCGTTCATCACTGACAAAGTATGCCAGTTGTCTGTGCATCCCCAGAC	4418
Db	4350	CTAAGCAGATCTCTCGCGTTCATCACTGACAAAGTATCCAGCTTGTCTGTGCATCCCCAAT	4409


```
OY 4419 TACATGTCACCGCCAGGACTGAATGTATACACTGGCTGGGAGAAACCCAGGTACC 4478
  || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4410 TATGTGTCGCTGACCGGACCGAATGTTTATCTACTGCTGGGAGAAACCCAGGTACT 4469
OY 4479 TTTGGAGCTGGCTCTTCTGAAGAAACCCAGCTCTGTTATGTAGAAATGATGTCCAA 4538
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4470 TTTGGAGCTGGCTCTTCTGAAGAAACCCAGCTCTGTTATGTAGAAATGATGTCCAA 4529
OY 4539 CACTATTAAGT-----ATATTTGTGCTGAGCTTTGGCC 4571
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4530 CCTTATGAGTTTCTGAATGAAGAGTCCATCCACCAACTCTGTCTGGCCATTGGCC 4589
OY 4572 AGAGCACTGACAGTTGGCCAGGCTGACAGTGGAGGCGCTGTGTTGCTTGAAGAGAC 4631
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4590 GGAGGCACTGACAGTTGGCCAGGCTGACAGTGGAGGCTGTGTTGCTTGAAGAGAGAC 4649
OY 4632 AATATCAATTTTACAGAGAGTCACTCTTGGGCTGTGGCTGCTGACGCCCCAATTAAGCT 4691
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4650 AATATCAATTTTACAGAGAGTCACTCTTGGGCTGTGGCTGCTGACGCCCCAATTAAGCT 4709
OY 4692 GGTGCTATGCTGCTGTTCAAGCTTGTACTTGTGATGAGGGAATGATGAGAAATTAAT 4751
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4710 GGTGCTATGCTGCTGTTCAAGCTTGTACTTGTGATGAGGGAATGATGAGAAATTAAT 4769
OY 4752 TATATGAGCGGAGGACAGAGTGAAGCATCAACCTTATGAAGCTGAACGTGGGTAAG 4811
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4770 TATATGAGCGGAGGACAGAGTGAAGCATCAACCTTATGAAGCTGAACGTGGGTAAG 4829
OY 4812 ATTTAGCATGCTGGAATATATAGACAGATCAACAGACAGTGTCCAGCTACAG 4871
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4830 ATTTAGCATGCTGGAATATATAGACAGATCAACAGACAGTGTCCAGCTACAG 4889
OY 4872 CTATGCGCAACCTTGCGACTTTTGGTATTTTGTGTATTAAGCTTGAAGCTGACTGAC 4931
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4890 CTATGCGCAACCTTGCGACTTTTGGTATTTTGTGTATTAAGCTTGAAGCTGACTGAC 4932
OY 4932 AATATCTATTAAGTGTACATAGCTATGACATTTGTTAAATTAACCTGCACTTAAT 4991
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4933 GATATCTGTAGTAAGTGCATACATAGCTATGACATTTGTTAAATTAACCTGCACTTAAC 4992
OY 4992 TT 4993
  || ||
Db 4993 TT 4994
```

RESULT 2
AAQ40319
ID AAQ40319 standard; cDNA; 2679 BP.
XX

AC AAQ40319;
XX
DT 17-AUG-1993 (first entry)
XX
DE Sequence of a plasminogen cDNA.
XX
KM Zymogen; fibrinolytic activity; cleavage; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 22..2454
FT /tag= a
XX
PN US5200340-A.
XX
PD 06-APR-1993.
XX
PF 22-MAY-1987; 87US-0053412.
XX
PR 22-MAY-1987; 87US-0053412.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX

```
PI Foster DC, Mulvihill ER, Ohara PJ, Pingel K, Yoshitake S;  
XX  
XX WPI: 1993-133739/16.  
DR P-PDB; AAR34428.  
XX  
XX Human tissue plasminogen activator single chain form fibrinolytic  
PT agent - comprises thrombin cleavable zymogen stimulating amide  
PT lytic activity, for lysing clots in heart attack and stroke  
PT victims and suppressing fibrin matrix  
PS  
PS Example; Fig 10A, 10B, 10C; 22pp; English.  
XX  
XX A lambda phage clone comprising a cDNA sequence encoding  
CC plasminogen was obtained from Dr. Mark Martzen at the University of  
CC Washington. The cDNA was isolated from a human liver library by  
CC probing with the partial sequence of Malinowski et al. The sequence  
CC of the complete cDNA and the encoded amino acid sequence are shown  
CC in AAQ40319 and AAR34428.  
XX  
SQ Sequence 2679 BP; 750 A; 659 C; 675 G; 595 T; 0 other;  
Query Match 24.7%; Score 1236.8; DB 14; Length 2679;  
Best local Similarity 81.1%; Pred. No. 0;  
Matches 1510; Conservative 0; Mismatches 272; Indels 80; Gaps 3;  
OY 3159 TATCGAGCATATTCCTCCACACTGTCTACAGAGAGACCTGTCAATCTTGGTATCTATG 3218  
  ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 868 TATCGGCGGAATGTGGCTGTATCCGTGTACCGTCCGGACACCTGTGACAGATGTGCAG 927  
OY 3219 ATACACACTGCGATCAGAGAGACCCAGAAACTACCCAAATGCTGGCTGACCGAGAAC 3278  
  || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 928 ACCCTTCACACATATACAGAGACACAGAAACTTCCCTTGCAAAATTTGGATGAGAAAC 987  
OY 3279 TACTGCAAGATCCAGATTTCTGGGAAACACCTGCTGTACACACCATGCTGTGTG 3338  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 988 TACTGCGGCATCTCTGACGAAAGAGGCGCCATGCTGCTACACACACAGCCAAAGTG 1047  
OY 3339 AGGAGGAGTACTGCAATGTGACACATCTCAGAAACAGATCAGTGTCTAGAGACT 3398  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1048 CGGTGGGAGTACTGTAAAGTATACCTGCTCTG----- 1076  
OY 3399 CCCACTGTGTTGCTCACTTCCAGCATGAGAGCTCACTTGAAGACACCACTGAGCAA 3458  
  || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1077 -----TGACTCTCTCCAGTATCCAGCAATTTGGTCCACAGCACTGAGACTA 1131  
OY 3459 ACCCTGTGCTCGGAGAGTACCATGTAATGSCAGAGTTATGAGGACATTCCTCC 3518  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1132 ACCCTGTGCTCGGAGAGTACCATGTAATGSCAGAGTTATGAGGACATTCCTCC 1191  
OY 3519 ACCACTGTACAGAGAAAGCATGTCAATCTTGTATCATGACACACACCGGATCAG 3578  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1192 ACCACACACAGAGAAAGAGTGTCAATCTTGTATGACACACACACCGGACAG 1251  
OY 3579 AGGACCCAGAGAAATACCAATGATGAGCTGACATGATCTCTCAGAAATCCAGAT 3638  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1252 AAGACCCAGAGAAATACCAATGATGAGCTGACATGATCTCTCAGAAATCCAGAT 1311  
OY 3639 GCCGATACAGGCGCTTGTGTATACATGACACCCACATCAGGTGGAGTACTGCAC 3698  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1312 ACCGATTAAGGCGCTGTGTATACACAGACCCACATCAGTGGAGTACTGTGAC 1371  
OY 3699 CTGACGAGATGCTCAGACACAGAGAGAGTGTGTCTCTGAGTGTATCAGAGTT 3758  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1372 CTGAAAAAATGCTCAGAGAAACAGAGAGTGTGTAGACCTCCGCTGTGTCTGTCT 1431  
OY 3759 CCAAGCTAGGCGCTCTCTCTGACACAGACTGTATGTTGGAAATGGAAGATACCGG 3818  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1432 CCAAGATGTAGAGACTCTCTCCAGAGAGACTGTATGTTGGAAATGGAAGATACCGG 1491  
OY 3819 GCGAAGAGCAACCATGTTACTGGGAGCCCATCCGAAATGGGCTGCCAGGAGCC 3878  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1492 GCGAAGAGGCGACACTGTACTGTGGAGCCCATCCAGAGACTGGGCTGCCAGGAGCC 1551
```

QY	3879	CATAGCACAGACGCTTATTCACGAGCACAATTAATGGCAGGCTCTGGAAAAAATTC	3933
Db	1552	CATAGCACAGGATTTTCTACCTACGAGACAAATCCACGCGGGCTCTGGAAAAAATTC	1611
QY	3939	TGCGCTAACCTCGAAGTGATCAATAAVGGTCCCTGGCTCTCACAAATGAATCCAGAAAA	3998
Db	1612	TGCGCTAACCTGATGGATGTAGGGGGTCCCTGGCTCTCACAAATGAATCCAGAAAA	1671
QY	3999	CTTTTGGACTACTGATATNCCCTCTCTGGCATCTCTTATTTGATTTGGGAAGCT	4058
Db	1672	CTTTGACTACTGATGATGCCCTCAGTGGGGCCCTCTATTTGATTTGGGAAGCT	1731
QY	4059	CAAGTGGACCCGAAATAATGTCTGGAAGCATTTGAAGGGGGTGTGGGCCCCACCAT	4118
Db	1732	CAAGTGGACCCGAAATAATGTCTGGAAGGGTGTGAAGGGGGTGTGGGCCCCACCAT	1791
QY	4119	TCTGGCCCTGGCAGTCTAGCTCTACAGAACAGTTTGGAAAGCACTCTGTGAGGCAAC	4178
Db	1792	TCTGGCCCTGGCAGTCTAGCTCTTGAACAAGTTTGGAAAGCACTCTGTGAGGCAAC	1851
QY	4179	TTAATATCCCGAGAGTGGGCTGTGACCTGTGCTCACTGCTTTGAAGAAGTCTCTAGGCT	4238
Db	1852	TTGATATCCCGAGAGTGGGCTTGTACGTGCTGCCACTGCTTGGAGAAGTCTCCAGAGCT	1911
QY	4239	TGATCTTACAAGGTACTCTGGGGGCACACCAAGAAAGAACTCGAATCTATGTGTCAG	4298
Db	1912	TGATCTTACAAGGTACTCTGGGGGCACACCAAGAAAGTATCTCGAACCGCATGTTCAG	1971
QY	4299	GAAATGGAAGTCTAGGCTGTTCTTGGAGCCACACAGACAGATATTTGCTTGTCTAAG	4358
Db	1972	GAAATGGAAGTCTAGGCTGTTCTTGGAGCCACACAGAAAGATATTTGCTTGTCTAAG	2033
QY	4359	CTAAGACGCGCTGCGTACTACAGCAAAAGTATGCGAGCTGTCTGCCATCCCGAC	4418
Db	2032	CTAAGACGCTGCGCGTACTACAGCAAAAGTATCCAGCTGTCTGCCATCCCGAAT	2091
QY	4419	TACATGTCACCGCCAGGACTGAATGTACACTGCGCTGGGGAGAAACCAAGGTAC	4478
Db	2092	TATGTGTGTGCTGACCCGAGCGAATGTTCATACACTGTGGCTGGGGAGAAACCAAGGTACT	2151
QY	4479	TTTGGGACGTGGGCTCTCAAGAAAGCCGACGCTCTTGTATTTAGATGAAGTGTCAAT	4538
Db	2152	TTTGGAGTGGGCTCTCTCAAGAAAGCCGACGCTCTGTGATTTGAGATTAAGTGTCAAT	2211
QY	4539	CACATTAAGT-----ATATTTGCTGAGCATTTGGCC	4571
Db	2212	CGGTATGAAATTTTCTGAATGAGAGATCCAACTCACCGAACTCTGTGCTGGGCACTTTGGGC	2271
QY	4572	AGAGGACATGACAGTTGGCAGAGGGTACAGTGGAGGGGCTCTGTGTTCTCTCGAAGAGAC	4633
Db	2272	GGAGGACATGACAGTTGGCAGAGGGTACAGAGTGGAGGGTCTGTGTTCTCTCGAAGAGAC	2333
QY	4633	AAATACATTTTACAGAGAGTCACTCTGTGGGGCTTTGGGCTGTGACGCCCCCAATAAGCT	4691
Db	2332	AAATACATTTTACAGAGAGTCACTCTGTGGGGCTTTGGGCTGTGACGCCCCCAATAAGCT	2391
QY	4692	GGAGTCTATGCTCGATTTCAAGTGTGTACTTGGATTGAGTGGGAATGATGAGAAATAT	4751
Db	2392	GGAGTCTATGCTCGATTTCAAGTGTGTACTTGGATTGAGTGGGAATGATGAGAAATAT	2451
QY	4752	TAAATTTGAGCGGGAGACAGATGAGCATCAACTTCTTGAAGCTGAAAGTGGGGTAAAG	4811
Db	2452	TAAATTTGAGCGGGAGACAGATGAGCATCAACTTCTTGAAGCTGAAAGTGGGGTAAAG	2511
QY	4812	ATTATGACATGCTGGAATTAATTAACAGCATCAAAAGAAAGCACTGTCTCCAGCTACAG	4871
Db	2512	ATTATGACATGCTGGAATTAATTAACAGCATCAAAAGAAAGCACTGTCTCCAGCTACAG	2571
QY	4872	CTATGCGCAAACTTGGCATTTTGTGTATTTTGTGTATTAAGCTTTTAAAGTCTGACTGAC	4931
Db	2572	CTATGCGCAAACTTGGCATTTTGTGTATTTTGTGTATTTTAAAGTCTGACTGACTAT	2614
QY	4932	AAATTTCTGATTAAGGTGTATAGCTATACATTTGTTAATAATAACTGTGACTATAT	4991

Dd	2615	GGATTCTGTAGTAAAGTCACATACATGACATTGTTAAAAATAA	CTGTACTTAAC	2674
Qy	4992	TT	4993	
Dd	2675	TT	2676	

```

RESULT 3
AAK35376
ID AAK35376 standard; DNM: 2732 BP.
XX
XX
AC AAK35376;
XX
DT 16-JUL-1999 (first entry)
XX
DE SEQ ID 51 of M0916889.
XX
KW Angiostatin; endostatin; interferon; thrombospondin;
interferon-inducible protein; platelet factor 4; anti-angiogenic;
anti-tumor; multifunctional protein; angiogenic-mediated disease;
cancer; diabetic retinopathy; macular degeneration; arthritis;
tumor cell production; ss.
XX
OS Homo sapiens.
XX
PN M0916889-A1.
XX
PD 08-APR-1999.
XX
XX
PF 30-SEP-1998; 98WO-US20464.
XX
PR 01-OCT-1997; 97US-0060609.
XX
XX
(PSEAR ) SEARLE & CO G D.
XX
PI Bojanowski MA, Caparon MH, Casperson GF, Gregory SA;
PI Klein BK, McKearn JP;
XX
XX
WR: 1999-255098/21.
XX
New multifunctional proteins useful for treating angiogenic-mediated
diseases
XX
PS Disclosure; Page 85-86; 121pp; English.
XX
XX
The specification describes multifunctional proteins which comprise
combinations of angiostatin, endostatin, interferon, thrombospondin,
interferon-inducible protein and platelet factor 4, and have
anti-angiogenic and/or anti-tumor activity. The multifunctional protein
may exhibit useful properties such as having similar or greater
biological activity when compared to a single factor or by having
improved half-life or decreased adverse side effects, or a combination
of these properties. The proteins can be used for treating an
angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
degeneration, or arthritis. They can also be used for inhibiting the
production of tumor cells (characteristic of lung, breast, ovarian,
prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
growth. The present sequence is used in the course of the invention.
XX
XX
Sequence 2732 BP; 757 A; 667 C; 690 G; 618 T; 0 other;
SQ
Query Match 24.7%; Score 1235.4; DB 20; Length 2732;
Best Local Similarity 80.9%; Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 276; Indels 80; Gaps
3;
3159 TATGAGGCAATCCCTCCACCACTGTCACAGAGAGAGACCTGTCATCTGGTCATCTATG 3218
|||||
901 TATGCGGGGAAATGCTGTATCCGCTGTCGCGGACACACTCTCTGACACTGAGTGCACAG 960
|||||
3219 ATACCACTACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3278
|||||

```

D	961	ACCCTCAGACATATACAGACACACAGAAAACCTCCCTCGAAAATTTTGGATGAAMAC	1020
Q	3279	TACTGACGAATCCAGATTCTGGGAAACAACCTTGGTGTACACACCGATCCGTGTG	3338
D	1021	TACTGCGCAATCTGACGGAAAAAGGCCCATGGTCCATCAACCAACGCAAAATG	1080
Q	3339	AGGTGGAGTACTCAATCGACACATGCTCGAAGACAGAAACAGATGCTCTAGAGACT	3398
D	1081	CGATGGAGTACTAGATACCGTCTG-----	1109
Q	3399	CCCACTGTGTTCAGTTCACAGCAGAGSGCTATTCAGACGACCACTAGCA	3458
D	1110	-----TGACTCTCCCGCATTCACAGGAACGTTGGCTCCACAGACACTGAGCTA	1164
Q	3459	ACCCTGTGTCCGGAGTGTACATGTGTAATGGCCAGATTAAGAGCAATTTCC	3518
D	1165	ACCCTGTGTCCAGAGTCTCAATGTGTGACAGACTACAGAGCAATCTCC	1224
Q	3519	ACCAGTCACAGAGAACATGTCAATTTGTATCCATGACACACACCGGATAG	3578
D	1225	ACCCACACACAGAAAGAAAGTGTCTGTGTATCTATGACACACCGGCACAG	1284
Q	3579	AGGACCCGAGAAACTACCCAAATGATGCGCTGACATGAACTATGCAAGATCCAGAT	3638
D	1285	AAGACCCGAGAAACTACCCAAATGCTGCGCTGACAAATGAATCTACGAGATCCAGAT	1344
Q	3639	GGCGATACAGGCGCTGTGTATTTACCAAGACCCGACACATAGTGGAGATACGAC	3698
D	1345	GGCGATACAGGCGCTGTGTATTTACCAAGACCCGACACATAGTGGAGATACGAC	1404
Q	3699	CTGACGCGATCTCAGACACAGAGGAGCTGTGTGCTCTCCGACTGTCAATGAGTT	3758
D	1405	CTGAAAAATCTCAGACACAGAGGAGTGTGTAGACACTCGCGCTGTGCTCGTT	1464
Q	3759	CCAGGCTAGGGCCCTCTGTGAACAAGACTGTATGTTGGAAATGGAAAGATACCG	3818
D	1465	CCAGATGATAGACTCTCTCGAAGAAAGCTGTATGTTGGAAATGGAAAGATACCGA	1524
Q	3819	GGCAGAGGCAACCACTGTACTGGGAGCCATCCAGAAATGGCTGCCAGAGGCC	3878
D	1525	GGCAGAGGAGGACCACTGTACTGGGAGCCATCCAGAACTGGCTGCCAGAGGCC	1584
Q	3879	CATAGCACACAGCTTCAATTCACAGGACCAATAATGGCAGGTCGCAAAAAATATAC	3938
D	1585	CATAGCACAGAGATTTCTACTCCAGAGCAATAATCAAGGGGCTCGAAAAAATATAC	1644
Q	3939	TGCGTAACCTGATGGTGACATCAATGTCCTCTGTCTACACATGAATCCAGAAAA	3998
D	1645	TGCGTAACCTGATGGTGATGATGATGTCCTCTGTCTACACAAATCCAGAAAA	1704
Q	3999	CTTTTGACTACTGTGATATCCCTCTGTGATCTCTTATTTGATTTGGGAAGCT	4058
D	1705	CTTTAGACTACTGTGATATCCCTCTGATGTGGGCCCTTATTTGATTTGGGAAGCT	1764
Q	4059	CAAGTGAAGCGAAGAAATGCTCGAAGCAATTTAGGGGGGTGTGGCCACCCACAT	4118
D	1765	CAAGTGAAGCGAAGAAATGCTCGAAGGTTGTGGGGGGGTGTGGGCCACCCACAT	1824
Q	4119	TCTGTGCGCTGGCAAGTCACTCAAAACAAGTTTGAAGACCTCTGTGGAGCAC	4178
D	1825	TCTGTGCGCTGGCAAGTCACTCAAAACAAGTTTGAAGTGCACCTCTGTGGAGCAC	1884
Q	4179	TTTATATCCCAAGTGGGTGTGACTGTCTCACTGCTTGAAGAAATCTCAGAGCT	4238
D	1885	TTTATATCCCAAGTGGGTGTGACTGTCTCACTGCTTGAAGAAATCTCAGAGGCT	1944
Q	4239	TGATCTCAAGGTCATCTGGGTGCACACCAAGAGTGAACCTCGAATCTATGTTCAG	4298
D	1945	TGATCTCAAGGTCATCTGGGTGCACACCAAGAGTGAATCTCGAATCGATGTTCAG	2004
Q	4299	GAATATGAAGTGTCTAGGCTGTCTTGGAGCCCAACAAGCAATATTCGCTTGAAG	4358
D	2005	GAATATGAAGTGTCTAGGCTGTCTTGGAGCCCAACAAGCAATATTCGCTTGAAG	2064

QY	4359	CTTAGCAGGCGCTCCCGCATCACTATGCAAGAAATGACCAAGCTTTGCTGCATCCCAAC	4418
Db	2065	CTAAGCAGCTCCTCCGCTCATCACTGACAAAGTAATCCACAGCTGTCTGCATCCCAAT	2124
QY	4419	TACATGCTCACCGCCACAGACTGAAATTTACATCACTGCTGGGGGAAACCCAGATACC	4478
Db	2125	TATGTGGTGCCTACCGGACCGCAATTTTTCATCAGCTGGCTGGGGGAAACCCAGATACC	2194
QY	4479	TTTGGAGCTCGCCCTCTCAAGAAAGCCAGCTCTTGTATTGAATAAGATGTGCAT	4538
Db	2185	TTTGGAGCTCGCCCTCTCAAGAAAGCCAGCTCCCTGTGATTGAGATAAAGTGTGCAT	2244
QY	4539	CACATATAGT-----AATTTGGCTGCAGCATTTGGCC	4571
Db	2245	CGCTATAGTTTCTGAATGGAGATGCCAATCCACCAACTCTGTGCTGGGATTTGGCC	2304
QY	4572	AGAGGCACTGACAGTTGGCCAGGGGTGACAGTGGAGGGCCCTGTGTTGGCTTGAGAAAGAC	4631
Db	2305	GGAGGCACTGACAGTTGGCCAGGGGTGACAGTGGAGGGCTCTGTGTTGGCTTGAGAAAGAC	2364
QY	4632	AAATACATTTTACAAAGAGTCACTCTGTGGGGCTTGGCTGACACGCCCAATTAAGCT	4691
Db	2365	AAATACATTTTACAAAGAGTCACTCTGTGGGGCTTGGCTGACACGCCCAATTAAGCT	2424
QY	4692	GGGTCTATGCTCGTGTTCACAGTTTGTACTTGGATTGAGGGAATGAGAAATAT	4751
Db	2425	GGGTCTATGCTCGTGTTCACAGTTTGTACTTGGATTGAGGGAATGAGAAATAT	2484
QY	4752	TAAATGACCGGAGACAGTGAAGATCAACCTACTTGAAGAGTGCAAACGCGGATAGG	4811
Db	2485	TAAATGACCGGAGACAGTGAAGATCAACCTACTTGAAGAGTGCAAACGCGGATAGG	2544
QY	4812	ATTTAGCATGCTGGAATTAATATACAGCAATCAAGAAAGACATCTGTCCAGCTACAG	4871
Db	2545	ATTTAGCATGCTGGAATTAATCTGCGAGTATCAAGAAAGAAACATCTGTCCAGCTACAG	2604
QY	4872	CTATGCCAAACCTTGGCATTTTGGTATTTTGTGTATTAAGCTTTAAAGTGTGACATAC	4931
Db	2605	CTATGCCAAACCTTGGCATTTTGGTATTTTGTGTATTT-----TTCTGACTGCT	2647
QY	4932	AAATTCGTGTTTAAGTGTATAGTATGATGATGATTTGTTAAATTAACCTGCACTTAT	4991
Db	2648	GGATTCCTGTAGTAAAGTACATATGCAATTTGTTTAAATTAACCTGCTACTTAAAC	2707
QY	4992	TTTGATTT 4998	
Db	2708	TTTGATTT 2714	
RESULT 4			
ABN81696	ID ABN81696 standard; DNA: 2732 BP.		
XX	ABN81696;		
XX	03-SEP-2002 (first entry)		
DE	Human plasminogen encoding DNA.		
XX			
KW	Human; plasminogen; lys plasminogen; glu plasminogen; A61; P22;		
KW	vascular endothelial; cell proliferation; anti-angiogenic; cancer;		
KW	rheumatoid arthritis; Crohn's disease; diabetic retinopathy; anti-cancer;		
KW	cytostatic; gynaecological; neuroprotective; antipsoriatic;		
KW	antihypertensive; dermatological; antineoplastic; antiallergic;		
KW	anti-inflammation; antidiabetic; ophthalmological; immunosuppressive;		
KW	cardiac; vulnery; vasotropic; anti-tumour; gene; ds.		
XX			
OS	Homo sapiens.		
XX			
FT	Key Location/Qualifiers		
FT	1..2487		
FT	CDS		
FT	/*tag= a		

FT /product= "plasminogen"
FT sig_peptide 55..111
FT /tag= b
FT mat_peptide 112..2484
FT /tag= c
FT /product= "glu plasminogen"
FT 343..2484
FT /tag= d
FT /product= "lys plasminogen"
FT 343..1515
FT mat_peptide /tag= e
FT /product= "A61 isoform"
FT 343..1524
FT /tag= f
FT /product= "A61 isoform"
FT 343..651
FT mat_peptide /tag= g
FT /product= "p22"
FT
PN WO20024328-A2.
XX
PD 06-JUN-2002.
XX
PE 28-NOV-2001; 2001WO-US44515.
XX
PR 28-NOV-2000; 2000US-253725P.
XX
PA (MAIS/) WAISMAN D M.
PA (KASS/) KASSAM G.
PA (KWON/) KWON M.
PI Waisman DM, Kassam G, Kwon M;
DR WPI: 2002-527706/56.
DR P-PSDB: ABB83795.
XX
PT Novel naturally occurring fragment A16 or P22 of plasminogen, useful
PT for treating for anti-angiogenic treatment of a mammal suffering from
PT cancer and inhibiting proliferation of vascular endothelial cells -
XX
PS Claim 26; Fig 2; 88pp; English.
XX
CC The invention relates to an isolated naturally occurring fragment A61
CC or P22 polypeptide (I) of plasminogen or an isolated polypeptide
CC comprising 103, 391 or 394 contiguous amino acids of amino terminal of
CC plasminogen (A883795). (I) has vascular endothelial cell proliferation
CC inhibitor activity. (I) is useful for anti-angiogenic treatment of a
CC mammal suffering from cancer, acoustic neuromas, neurofibromas,
CC trachomas, pyogenic granulomas, telangiectasias, psoriasis, scleroderma,
CC atherosclerosis, rheumatoid arthritis, Crohn's disease, endometriosis,
CC adipsity, pyogenic granuloma, rubosis, and diabetic retinopathy,
CC retinopathy of prematurity, neovascular glaucoma, retrolental
CC fibroplasia, graft rejection, myocardial angiodenosis, plaque
CC neovascularization, haemophilic joints, angiodioma and wound
CC granulation and inhibiting proliferation of vascular endothelial cells,
CC in combination with an anti-cancer agent e.g. methotrexate, mitozantrene,
CC paclitaxel, vinblastine, 5-fluorouracil, cisplatin, leucovorin,
CC cyclophosphamide and oncolytic virus.
XX
SQ Sequence 2732 BP; 758 A; 669 C; 688 G; 617 T; 0 other:

Query Match 24.7%: Score 1235.4; DB 24; Length 2732;
Best Local Similarity 80.9%: Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 276; Indels 80; Gaps 3;

QY 3159 TATCGAGCATATCCCTCACACCATGTCAGAGAGACCTGTAATCTTGTCATCTATNG 3218
DB 901 TATCGCGGAGANTGGCGTTACGCTTCCGGGACACACCTGTAGACACTGGAGTGCACAG 960
QY 3219 ATACCACTGCGATCAGAGAGACCCAGAAACTACCCAAATGCTGGCCTGACCGAGAAC 3278
DB 961 ACCCTCACACATACAGAGACACCAAAACTCCCTGCAAAAATTTGGATGAAAC 1020

QY 3279 TACTGCAGAAATCCAGATTTCTGGAAACAAACCCCTGTTTACACACCGATCCGTGTG 3338
DB 1021 TACTGCCGCAATCTCTGACGGAAAAAGGCCCTATGTGCTTACACACCAACAGCAAGT 1080
QY 3339 AGGTGGAGTACTGCAATCTGACACAAATGCTCAGAAACGAATGAGTGTCTAGACACT 3398
DB 1081 CGGTGGGAGTACTGTAAGTACCGTCTG----- 1109
QY 3399 CCCACTGTTGTCCAGTTCACAGCATGGAGGCTCATCTGAGAGAGACCAACTGACCA 3458
DB 1110 -----TGACTCTCCCACTATCCAGAACCAATTGGTCCACAGACCACTGAGCTA 1164
QY 3459 ACCCTGTGGTCCGAGTGTCTACCATGGTAATGGCCAGAGTTATCGAGGACATTTCTCC 3518
DB 1165 ACCCTGTGGTCCAGAGACTACCATGGTATGGACAGACACTACAGGACACATCTCTCC 1224
QY 3519 ACCACTGTACAGAGAGCATGTCAATCTGTGATCTGATGACACACACACCGGCAATCAG 3578
DB 1225 ACCACACACAGAGAAAGATGTGATGTGTATCTATGACACACACACCGGACACAG 1284
QY 3579 AGGACCCAGAAACTACCCAAATGATGSCCTGACAACTACTCTCAGAGAAATCCAGAT 3638
DB 1285 AAGACCCAGAAACTACCCAAATGCTGSCCTGACAACTACTCTCAGAGAAATCCAGAT 1344
QY 3639 GCCGATAGAGGCCCTGTGTGTTTACCATGAGACCCACATCAGTGGAGTACTCTGAC 3698
DB 1345 GCCGATAGAGGCCCTGTGTGTTTACCATGAGACCCACGAGTGGAGTACTCTGAC 1404
QY 3699 CTGACGCGATGCTCAGACACAGAGAGAGCTGTGCTCTCTCCGACTGTATCCAGGTT 3758
DB 1405 CTGAAAAAATGCTCAGAGACACAGAGAGAGTGTGAGACCTCCGCTGTGTCTGCTT 1464
QY 3759 CCAAGCTAGGGGCTCTCTGACACAGCTGTATGTTGGAGATGGGAAAGATACCGG 3818
DB 1465 CCGATGTAGAGACTCTCTCCAGAAAGCTGTATGTTGGAGATGGGAAAGATACCGG 1524
QY 3819 GGCAGAGAGCAACACTGTTACTGGAGGCCATGCCAGAAATGGCTGCCAGAGGCC 3878
DB 1525 GGCAGAGAGGCGCACCTGTACTGGAGGCCATGCCAGAGTGGCTGCCAGAGGCC 1584
QY 3879 CATAGACACAGCAGTTCATTCAGAGCAAAATAAGGGGAGGCTGGGAAAAATTCAC 3938
DB 1585 CATAGACACAGCAGTTCATTCATCAGAGCAAAATCCAGGGGGGTGGGAAAAATTCAC 1644
QY 3939 TGGCGTAACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3998
DB 1645 TGGCGTAACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1704
QY 3999 CTTTGTGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4058
DB 1705 CTTTGTGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1764
QY 4059 CAAGTGAGCCAGAAAGATGCTGGAAGCATGTTAGGGGGGTGTGTGCCACCCACAT 4118
DB 1765 CAAGTGAGCCAGAAAGATGCTGGAAGCATGTTAGGGGGGTGTGTGCCACCCACAT 1824
QY 4119 TCGTGCCCTGGCAAGTCACTGTCAGACAGAGTTTGGAAAGCACTCTGTGAGGCACC 4178
DB 1825 TCGTGCCCTGGCAAGTCACTGTCAGACAGAGTTTGGAAAGCACTCTGTGAGGCACC 1884
QY 4179 TTAAATVCCCAAGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4238
DB 1885 TTGAATVCCCAAGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
QY 4239 TCAATCTTCAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4298
DB 1945 TCAATCTTCAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2004
QY 4299 GAAATAGAGTGTCTAGAGCTGTTCTTGGAGCCACACAGACAGATATGCTGCTTAAAG 4358
DB 2005 GAAATAGAGTGTCTAGAGCTGTTCTTGGAGCCACACAGAAAGATATGCTGCTTAAAG 2064
QY 4359 CTAGAGAGGCTGCGGTATCATCATGACAAAGTATGTCGACACTGTCTGCATCCCAAGAC 4418

```
Db 2065 CTAAGCAGTCCGTCGCTCATCTGACAAAGTATCCACGCTGTCTGCAATCCCAAT 2124
OY 4419 TACATGGTACCCGCGAGACTGAATGTACATCATTGCTGGGAGAAACCCAGATACC 4478
Db 2125 TATGTGGTCCGTGACCGGACCGAATGTTCATCAGCTGGGAGAAACCCAGATACC 2184
OY 4479 TTTGGAGCTGGCTTCTCAAGAAAGCCAGACTCTGTATTGAGAAATGAATGTGCAT 4538
Db 2185 TTTGGAGCTGGCTTCTCAAGAAAGCCAGACTCTGTATTGAGAAATGAATGTGCAT 2244
OY 4539 CACTATAGT-----ATAATTGGCTGAGCATTTGGCC 4571
Db 2245 CGATATAGATTTCGAATGGAAGATCCATCCACCGCAATCTGTGCTGGCATTTGGCC 2304
OY 4572 AGAGGCACTGACAGTGGCCAGAGGTGACAGTGAAGGCTCTGATTGCTTCGAGAAGAC 4631
Db 2305 GGAGGCACTGACAGTGGCCAGAGGTGACAGTGAAGGCTCTGATTGCTTCGAGAAGAC 2364
OY 4632 AAATACATTTTACAGAGATCTCTTGGGCTCTTGGCTGTGACGCCCCAATAAGCT 4691
Db 2365 AAATACATTTTACAGAGATCTCTTGGGCTCTTGGCTGTGACGCCCCAATAAGCT 2424
OY 4692 GGTGTATGCTGCTGTTTCAAGTTCATGATTCGATTCGAGGAATGATGAATAAT 4751
Db 2425 GGTGTATGCTGCTGTTTCAAGTTCATGATTCGATTCGAGGAATGATGAATAAT 2484
OY 4752 TAAATGACGGGAGACAGATGAACATCACTACTTGAAGCTGAAGCTGGGTAAGG 4811
Db 2485 TAAATGACGGGAGACAGATGAACATCACTACTTGAAGCTGGGTAAGG 2544
OY 4812 ATTTAGCATGCTGGAATTAATAGACATCAACAGAACACTGTTCCAGTACCAG 4871
Db 2545 ATTTAGCATGCTGGAATTAATAGACATCAACAGAACACTGTTCCAGTACCAG 2604
OY 4872 CTAGCCCAACCTTGGCATTTTGGTATTTTGTATTAAGCTTTAAGCTGTGACTGAC 4931
Db 2605 CTAGCCCAACCTTGGCATTTTGGTATTTTGTATTAAGCTTTAAGCTGTGACTGAC 2647
OY 4932 AAATCTGATTAAAGTGTATAGATCATGATTCATTTGTAATAAATAAAGCTGACTTAT 4991
Db 2648 GGATTCGTAGTAAAGTGAATGACATGACTATGACATTTGTAATAAATAAAGCTGACTTAC 2707
OY 4992 TTGATTT 4998
Db 2708 TTGATTT 2714

RESULT 5
AAQ11998
ID AAQ11998 standard; DNA: 6010 BP.
XX
AC AAQ11998;
XX
DT 05-SEP-1991 (first entry)
XX
DE Human plasminogen with substd. Arg561 from pUC119PM127.6 vector.
XX
KM Plasminogen; proteolytic cleavage; variant; thrombosis; HPG; ds.
XX
OS Homo sapiens.
XX
FH Key CDS 2329..4758
FT location/Qualifiers
FT 2329..4758
FT sig_peptide
FT mat_peptide 2386..4758
FT misc_recomb 2284
FT /*tag= d
FT /*tag= "end pUC119 vector"
```

```
FT misc_recomb 2285
FT /*tag= e
FT /note= "pmgn cDNA linker start"
FT 5'UTR 2294
FT /*tag= f
FT /note= "pmgn mRNA 5'UT start"
FT misc_RNA 2635..2874
FT /*tag= g
FT /label= kring1e_1
FT misc_RNA 2881..3117
FT /*tag= h
FT /label= kring1e_2
FT misc_RNA 3151..3387
FT /*tag= i
FT /label= kring1e_3
FT misc_RNA 3457..3693
FT /*tag= j
FT /label= kring1e_4
FT misc_RNA 3769..4011
FT /*tag= k
FT /label= kring1e_5
FT misc_RNA 4009
FT /*tag= l
FT /note= "start pmgn protein protease"
FT 3'UTR 4759
FT /*tag= m
FT /note= "start pmgn mRNA 3'UT"
FT polyA_site 4986..5074
FT /*tag= n
FT /label= poly(A)+
FT misc_recomb 5087
FT /*tag= o
FT /note= "start pUC119 vector and end cDNA 127 linker"
FT misc_feature 4066..4068
FT /*tag= p
FT /label= mutation site
FT /note= "wild-type: protease activation Arg"

WO9108297-A.
XX
PN 13-JUN-1991.
XX
PD 31-OCT-1990; 90WO-US06345.
XX
PF 01-DEC-1989; 89US-0444584.
XX
PR (GENE ) GENENTECH INC.
XX
PA Castellino FJ, Higgins D L;
XX
PI WPI, 1991-193201/26.
XX
DR P-PSDB; AARI2406, AARI3219, AARI3220.
XX
PT New human plasminogen variant with replaced ARG-561 - is
PT complexed with fibrinolytic enzyme for use in thrombolytic
PT therapy
XX
PS Disclosure; Fig 1(A-P); 82pp; English.
XX
XX The nucleotide at position 3809 is T for the vector pUC119PM127.6,
XX whereas Fig 1 of the specification reflecting the sequence
XX encoding native human plasminogen, contains a C. The codons at
XX positions 4066-4068 (NNN) which encode the amino acid Arg561 in the
XX wild-type plasminogen, are substituted to encode any amino acid
XX except Lys. Pref. the codon encodes amino acids Ser, Gly or Glu.
XX Position 561 is the critical cleavage site in the conversion of
XX plasminogen to plasmin. The resulting product is proteolytic
XX resistant plasminogen which may be used to treat thrombosis in
XX humans. When complexed with streptokinase, it does not degrade into
XX plasmin and it caused rapid plasminogen activator activity.
XX
SQ Sequence 6010 BP; 1611 A; 1497 C; 1475 G; 1424 T; 3 other;
```


OS	Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
XX	CDS	22..2454
FT		/*tag= a
FT		/transl_except= (pos 199..201, aa: Tyr)
FT		/transl_except= (pos 592..594, aa: Gly)
FT		/transl_except= (pos 688..690, aa: Gly)
FT		/transl_except= (pos 859..861, aa: Gly)
FT		/transl_except= (pos 892..894, aa: Val)
FT		/transl_except= (pos 1528..1530, aa: Glu)
FT		/transl_except= (pos 1840..1842, aa: Cys)
FT		/transl_except= (pos 1864..1866, aa: Glu)
FT		/transl_except= (pos 1993..1995, aa: Phe)
FT		/transl_except= (pos 2146..2148, aa: Gly)
FT	unsure	/note= "encodes protein plasminogen"
FT		/note= "1206
FT		/tag= b
FT		/note= "encodes Gly"
XX		
XX	US5648254-A.	
XX		
XX	15-JUL-1997.	
XX		
XX	14-JUL-1994;	94US-0275076.
XX		
XX	04-DEC-1989;	89US-0445302.
XX	15-JAN-1988;	88US-0144357.
XX	28-OCT-1991;	91US-0785865.
XX	14-JUL-1994;	94US-0275076.
XX		
XX	(ZYMO) ZYMOGENETICS INC.	
XX		
XX	Kumar AA, Mulvihill ER;	
XX		
XX	WPI: 1997-372063/34.	
XX	P-PSDB: AAM31169.	
XX		
XX	Production of recombinant plasminogen - by co-expression with	
XX	plasminogen-processing or -stabilising protein	
XX		
XX	Example 2; Fig 6A-D; 32pp: English.	
XX		
XX	This cDNA encoding for plasminogen was isolated from a lambda phage	
XX	library screened with a partial cDNA clone. This cDNA is used in a new	
XX	process for the production of plasminogen where a first DNA sequence	
XX	encoding plasminogen and at least one additional DNA sequence encoding a	
XX	protein that processes or stabilises the plasminogen is introduced into	
XX	a eukaryotic host cell. The protein is selected from alpha-1-antitrypsin	
XX	(AAT) and its variants and Argseirpsin. Both the DNA sequences are	
XX	operably linked to transcriptional promoter and terminator sequences. The	
XX	host cell is cultured under conditions that allow the DNA sequences to be	
XX	expressed and the recombinant plasminogen is isolated from the host cell.	
XX	Co-expression of plasminogen and the protein gives increased yields of	
XX	ungraded plasminogen.	
XX		
XX	Sequence 2679 BP; 751 A; 661 C; 670 G; 596 T; 1 other;	
XX		
XX	Query Match	24.5%; Score 1226.8; DB 18; Length 2679;
XX	Best Local Similarity	80.7%; Pred. No. 0;
XX	Matches 1503; Conservative	1; Mismatches 278; Indels 80; Gaps 3;
XX		
XX	3159 TATCGAGGCAATTCCTCCACCACTGTACAGGAGAGACCTGTCAATCTGGTCATCTANG	3218
XX		
XX	868 TATCGTGGGATGTGGCTGTACCTCTCCGGGACACCTGTGACACTGGAGTGCAG	927
XX		
XX	3219 ATACCACTACGGGATGAGAGACCCAGAAAATCTCAAAATGTGCTGGCCAGACGAGAC	3278
XX		
XX	928 ACCCTTACCACTATACAGACACCAAAAATCTTCCCTCAAAAATTTTGATGAAAC	987
XX		
XX	3279 TACTGCAGGAATCCAGATTCTGGGAAACAACCTGGTGTTCACACACCGATCCGTGTG	3338
XX		
XX	988 TACTGCAGGAATCCAGAAAAGGGCCCATGTGCTCCATACACCAACAGCAAGTG	1047

QY	3339	AGGTGGAGTACTGCATCTGACACAACTGCTCGAAGAACAGATCAAGTGTCTCTAGAGACT	3398
Db	1048	CGGTGGAGTACTGTAGATACCGTCTCTG-----	1076
QY	3399	CCCACTGTTGTTCCAGTTCACAGCATGAGAGGCTCATCTTGAAAGCAGACCAACTAGACA	3458
Db	1077	-----TGACTCTCTCCCAAGTATCCACGAAACAAATGTCGCCACAGCACACCTACGTGTA	1131
QY	3459	ACCCTGTGTGCCGAGTGTACTACATGATTAAGGGCAGAGTTATGAGAGCACAATCTTCC	3518
Db	1132	ACCCTGTGTGCCAGACTGTATACATGATGATGAGACAGAGGCTACCGAGAGCACAATCTTCC	1191
QY	3519	ACCCTGTACAGAGAGACATGTCAATCTTGGTCACTCATGACACACACCGCATAG	3578
Db	1192	ACAAACCAACAGSAAAGAGTGCATCTTGGTCATCTATGACACACACCGGCACAG	1251
QY	3579	AGGACCCCAAGAAACTCCCAAAATGATGGCTACAAATGACATCACTGACGAGAAATCCAGT	3638
Db	1252	AAGACCCCAAGAAACTCCCAAAATGCTGGCTTACATCACTACTCTGACGAGAAATCCAGT	1311
QY	3639	GCCGATACAGGCCCTTGTTGTTTACCATGAGCCACAGCATCAGTGGAGTACTGCAAC	3698
Db	1312	GCCGATTAAGGCCCTTGTTTATCCACAGACCCACAGTCAAGTGGAGTACTGCAAC	1371
QY	3699	CTGACGCATCTCTAGACACAGAAAGSACGTGGTGCCTCCGACTCTCATCCAGTT	3758
Db	1372	CTGAAAAATCTCTAGAGAACAGAAAGCGATGTTGTAGCACCTCCGCTGTCTCTGCTT	1431
QY	3759	CCAAAGCTTAGGGCCCTCTCTGCAACAAGATGTATGTTTGGAAATGGAAAGATACGG	3818
Db	1432	CCAGATGTAGAGACTCTCTCGAAGAAAGACTGTATGTTGGAAATGGAAAGATACCGA	1491
QY	3819	GGCAAGAAAGCAACACTGTTTACTGGGAGCCCATCCAGAAATGGGCTCCACAGAGCC	3878
Db	1492	GGCAAGAGGGGACCACTGTTTACTGGGAGCCCATCCAGAACTGGGCTCCACAGAGCC	1551
QY	3879	CATAGACACACAGCTTCATCCAGGACAAATTAATGGGACAGTCTGGAATAAAATTAAC	3938
Db	1552	CATAGACACACACTTTTCACTCCAGACAAACAAATCCACGGCGGTGTGGAAAAATTAAC	1611
QY	3939	TGCCGTAAACCTGATGGTGCATCATATGTTCCCTGCTCAACAATGAATCAAGAAA	3998
Db	1612	TGCCGTAAACCTGATGGTGAATGTAGGTGGTCCCTGGTGTACACGCAAAATCAAGAAA	1671
QY	3999	CTTTTGTACTCTGATATCCCTCTCTGTCACCTCTTCAATGATTTGGGAAAGCT	4058
Db	1672	CTTTTGTACTCTGATATCCCTCTCTGTCAGCTCTTCAATGATTTGGGAAAGCT	1731
QY	4059	CAAGTGAAGCCGAAGAAATGCTGGAAGACTTGTAGGGGGGTGTGTGGCCCAACCCACAT	4118
Db	1732	CAAGTGAAGCCGAAGAAATGCTGGAAGGGTGTAGGGGGGTGTGTGGCCCAACCCACAT	1791
QY	4119	TCTGTGGCCCTGGCAAGTCAGTCTCAAGAAAGGTTTGGAAAGACTTCTGTGGAGCAC	4178
Db	1792	TCTGTGGCCCTGGCAAGTCAGCTTGAAGAAAGGTTTGGAAATGCACTTCTGTGGAGCAC	1851
QY	4179	TTAATATCCCAAGATGGGTGCTGACTGCTGCTACATGCTTGAAGAAAGCTCAAGAGCT	4238
Db	1852	TTGATATCCCAAGATGGGTGTGACGTGCTGCCACATGCTTGGAGAAAGTCCCAAGGCT	1911
QY	4239	TGATCTCAAGAGTCATCTGTGGTGCACACCAAGAAAGTGAACCTGATCTCATTTTAC	4298
Db	1912	TGATCTCAAGAGTCATCTGTGGTGCACACCAAGAAAGTGAATCTGCAACCGCATTTTAC	1971
QY	4299	GAAATAGAAAGTGTAGAGCTTCTTGTGGACCCACACAGCAAGATTTTCTCTGTATAAG	4358
Db	1972	GAAATAGAAAGTGTAGAGCTTCTTGTGGACCCACACAGAAAGATTTTCTCTGTATAAG	2031
QY	4359	CTAAGCAGGCTGCGTGTATCACTGACAAAGTAATGCCAGCTTGTCTGCAATCCCAAGAC	4418
Db	2032	CTAAGCAGTCTCTGCGTATCACTAGCAAAAGTAATCCAGCTTGTCTGCAATCCCAAT	2091

QY	4419	TACATGGCACCGGCAGACTGATGATTTACATGACAGCGTGGGGAGAAACCCAAAGTACC	4478
Db	2092	TATGTGGTCGTGACCGGACCGAATGTTTCATCTACGTGGTGGGAGAAACCCAACTACT	2157
QY	4479	TTTGGGACTGGCCTTCTTCAGAAAGCCCAAGCTCCTGTTATTGAGAAATGAAGTGCANT	4538
Db	2152	TTTTTGGACTGGCCTTCTTCAGAAAGCCCAAGCTCCTGTTATTGAGAAATGAAGTGCANT	2211
QY	4539	CACATATAGT-----ATATTGTCGTAGCAATTTGGCC	4571
Db	2212	CGCATATGAGTTCGTGAATGGAAGAATGCCAATCCACCGAACTCTGTCTTGCGATTTTGGCC	2271
QY	4572	AGAGGCACTGACAGTTTGCCAGGGTGCACATGGAGAGGCTCTGTGTTGCTTCGAGAAAGAC	4631
Db	2272	GGAGGCACTGACAGTTTGCCAGGGTGCACATGGAGAGGCTCTGTGTTGCTTCGAGAAAGAC	2331
QY	4632	AAATACATTTTACAMAGAGTCACTCTTTGGGGCTTGCGCTGCACGCCCCAAATTAAGCT	4691
Db	2332	AAATACATTTTACAMAGAGTCACTCTTTGGGGCTTGCGCTGCACGCCCCAAATTAAGCT	2391
QY	4692	GGTGTCAATCTGCTGTTTCAAGGTTTGTTACTTGATTTGAGAGGAATGATGAAATAT	4751
Db	2392	GGTGTCAATCTGCTGTTTCAAGGTTTGTTACTTGATTTGAGAGGAATGATGAAATAT	2451
QY	4752	TAAATGGACGGGAGACAGAGTGAAGCATCAACCTACTTATGAAGCTGAAACGTGGGTAGG	4811
Db	2452	TAAATGGACGGGAGACAGAGTGAAGCATCAACCTACTTATGAAGCTGAAACGTGGGTAGG	2511
QY	4812	ATTAGAGATCTGGSAAATATATAGACAGCAATCAAGAGACACATGTTCCAGCTAACGG	4871
Db	2512	ATTAGAGATCTGGSAAATATATAGAGTAATCACTGGCAGATATCAAGAGACACATGTTCCAGCTAACGG	2571
QY	4872	CTATGCCAAACCTGGCATTTTTGGTATTTTGTGTATTAAGCTTTTAAGTCTGACTGAC	4931
Db	2572	CTATGCCAAACCTGGCATTTTTGGTATTTTGTGTATTTTGTGTATTTTGTGTATTTTGTGTATTT	2631
QY	4932	AAATTCGTATTAAGGCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT	4991
Db	2632	GGATTCGTATTAAGGCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2691
QY	4992	TT 4993	
Db	2692	TT 2696	
RESULT 7			
ID	AA012547		
	AA012547 standard; DNA: 2753 BP.		
XX	AA012547:		
AC			
XX	23-SEP-1991 (first entry)		
DT			
XX	Encodes Plasminogen mutEIN T1 with thrombin cleavage site.		
DE			
XX	protease; fibrinolysis; blood clotting; ss.		
KW			
XX	Key	Location/Qualifiers	
FH	mutation	1796..1801	
FT		/*tag= a	
FT		/note= "replaces COTGGA (Pro-Gly) with GGTCT	
FT		(Gly-Pro)"	
FT	CDS	65..2497	
FT		/*tag= b	
FT		/product= modified plasminogen	
XX	W09109118-A.		
XX	27-JUN-1991.		
PD			
XX	07-DEC-1990.	90WO-G001912.	
PF			
XX	07-DEC-1989.	89GB-0027722.	
RR			

Pr	07-DEC-1990;		90WO-G8B1911.	
Px	(BRBT)-		BRIT BIO-TECHN LTD.	
Ft	Dawson KM,		Edwards RM,	Forman JM;
Dr	MP1; 1991-208145/28.			
Xx	P-P-SDB; AAR12938.			
CC	Activatable fibrinolytic and antifibrinolytic proteins - activated by e.g. factor Xa, thrombin or activated protein C			
PS	Claim 16; Fig 2 and Fig 5; 73pp: English.			
XX	This sequence encodes a plasminogen mutant cleavable by thrombin. Activation is localised to the thrombus because cleavage to plasmin is by an enzyme of the blood clotting pathway. Compositions comprising the mutant plasminogen encoded by this sequence are used for treatment or prevention of thrombosis, etc.			
CC	See AA012542-Q12558.			
SQ	Sequence 2753 BP; 764 A; 671 C; 695 G; 623 T; 0 other;			
Query Match		24.5%; Score 1224.2; DB 12; Length 2753;		
Best Local Similarity		80.6%; Pred. No. 0;		
Matches 1504; Conservative		0; Mismatches 283; Indels 80; Gaps		
Qy	3159	TATTCGAGCATTCCTCCACACGTCACAGAAGAGACTTGTCAATTGGTGCATCATTG	3218	
Db	911	TATTCGGGGGAATGTGGCGTTTACCGGTGCCGGCACACCTGTCAAGCTGAAGTGCACAG	970	
Qy	3219	ATVACACACTGGCATCAGAGAGACCAGAAAACCTACCCAAATGCTGGCCTGACCCGANAAC	3278	
Db	971	ACCCTTCACACACATATAAGAGACACAGAAAACTTCCCTGCAAAATTTGGATGAAAAC	10330	
Qy	3279	TACTGAGGAATCCGATTCGGGAAACAACCTGTGTATACACACGATCCGCTGTG	3338	
Db	1031	TACTGCCGAAATCTCGACAGGAAAAAGGCCCATGTGTGCCTATACACCAAGCCAAGTG	10990	
Qy	3339	AGGTGGAGTACTGCATCTGCACATCTGCACATCTGAGAAACGAAATCGATGTCTTAGAGACT	3398	
Db	1091	CGGTGGAGTACTGTAAATACGCTCTG-----	1119	
Qy	3399	CCCACGTTTTCACAGTTTCCAAAGATGAGAGGCTCATFTGAAAGAGCACCACATGACGA	3458	
Db	1120	-----TGACTCTCTCCACAGTATCCAGGAAACAAATTGGCTCCACAGACACACTAGAGTA	1174	
Qy	3459	ACCCCTGTGTCCGGCAGTGTACACATGTAATGGCCAGAGTATATGAGCGACATTTCTCC	3518	
Db	1175	ACCCCTGTGTGTCCAGAGATGTGTACATGTGTATGAGACAGACCTACCGAGGACATCTTCC	1234	
Qy	3519	ACCACGTGCACAGNAAGACATGTGCATCTGGGTATCCATGACACACACCGGCATCG	3578	
Db	1235	ACACACACACAGAAAAGATGTGCAAGTCTTGCTATATGACACACACCGCGCACAG	1294	
Qy	3579	AGGACCCCAAGAAATACCCCAATGATGGCTGTACATGAATGAATCTACAGAAATCCAGAT	3638	
Db	1295	AAGACCCCAAGAAATATACCAAATGTGTGCTGTGACAAATGAATCTACTCAGAGATCCAGAT	1354	
Qy	3639	GCCGATACAGGCGCTGTGGTGTATTTACCATGAGAACCCCAATCATCAGTGGGAGTCTGCAAC	3698	
Db	1355	GCCGATAAAGAGCCCTGTGTGTATTTACCAACAGACCCCAATCATCAGTGGGAGTGTGCAAC	1414	
Qy	3699	CTGACGCGATGTCTGACACACAGAAGAGGACTGTGTGGCTCTCCGACTGTCAATCCAGTT	3758	
Db	1415	CTGAAAAAATGTCTGAGAACAGAAAGAGAGATGTGTGACACCTCCGCGCTGTTCCTGCTT	1474	
Qy	3759	CCAAAGCTTAGGGCCCTCTCTGAAACAAGACTATGTTTGGGAATGGGAAGGATACCGG	3818	
Db	1475	CCAGATGTAGAGATCTCTTCCGAAAGAACTGTATGTTTGGGAATGGGAAGGATRACGA	1534	
Qy	3819	GCCAAAGCAACCACTGTTTACTTGGGAGCGCATSCCAGAAATGGGCTGCCAGAGACC	3878	


```
Db 1535 GGCAGAGGGGCGACCTGTTACTGGAGCCCATGCCAGAGCTGGGCTGCCAGAGCCC 1594
OY 3879 CATAGACACAGCAGTTTCATTTCCAGGAGCAAAATMAATGGGCAAGTCTGAAAAAAATTAC 3938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1595 CATAGACACAGCATTTTTCACGTCCAGAGCAAAATCCAGGGGGGGTCTGAAAAAAATTAC 1654
OY 3939 TGGCGTAACCTGATGCTGACATCATATGTCCTGGTGTCTACCAATGAATCCAGAAAA 3998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1655 TGGCGTAACCTGATGCTGATGAGTGTCTGGTGTCTACCAATGAATCCAGAAAA 1714
OY 3999 CTTTTCGACTACTGTGATATCCCTCTCTGTGACATCCCTTCATTTGATTTGGGAAGCCT 4058
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1715 CTTTTCGACTACTGTGATATCCCTCTCTGTGACATCCCTTCATTTGATTTGGGAAGCCT 1774
OY 4059 CAAGTGAAGCCGAGAAATATGCTCGAAGCATTTGAGGGGGGTGTGTGAGCCCAAT 4118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1775 CAAGTGAAGCCGAGAAATATGCTCGAAGCATTTGAGGGGGGTGTGTGAGCCCAAT 1834
OY 4119 TCCCTGGCCCTGGCAAGTCAGTCAGAAAGTTGGAAAGCATTCCTGTGGAGGACAC 4178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1835 TCCCTGGCCCTGGCAAGTCAGTCAGAAAGTTGGAAAGCATTCCTGTGGAGGACAC 1894
OY 4179 TTAATATCCCAAGAGTGGGTGCTGACTGCTCATCTGCTTGAAGAAGTCTCAAGGCT 4238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1895 TTGATATCCCAAGAGTGGGTGCTGACTGCTGACACAGTGTGAGAAAGTCTCAAGGCT 1954
OY 4239 TCATCCCTAAGATCATCCCGGGTGCACACCAAGAGTGAACCTCAATCTCATGTTTCAG 4298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1955 TCATCCCTAAGATCATCCCGGGTGCACACCAAGAGTGAACCTCAATCCAGGACGCTCAG 2014
OY 4299 GAAATGAAGTGTCTAGGCTGTCTTGAGGCCACACACAAAGCATATTCCTTGTCTAAG 4358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2015 GAAATGAAGATGTCTAGGCTGTCTTGAGGCCACACAAAGATATTTGCTTGTCTAAG 2074
OY 4359 CTAAGCAGGCTGCCGCTCATCTGACAAAGTAATGTCAGCTGTGTGCCATCCCGAGAC 4418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2075 CTAAGCAGTCTGCCGCTCATCTGACAAAGTAATCCAGCTGTGTGCCATCCCGAAAT 2134
OY 4419 TACATGCTCACCGCCAGAGCTGAATGTACATCAGTGGTGGGAGAAAAACCAAGTACC 4478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2135 TATGTCGTGCTGACCGGACCGAATGTTTATCATCTGCTGGGAGAAACCAAGTATCT 2194
OY 4479 TTTGGAGCTGGCTTCTCAGAGAACCCAGCTCTTGTATTGAAGATGAATGTCCAAAT 4538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2195 TTTGGAGCTGGCTTCTCAGAGAACCCAGCTCTTGTGTGAATGAATGAATGTCCAAAT 2254
OY 4539 CACTATAGT-----ATATTGTGCTGACATTTGGCC 4571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2255 CCCTATGATTTTGAATGGAAGAGTCCAAATCCACGAATCTGTGCTGGGCAATTTGGCC 2314
OY 4572 AGAGGCACTGACAGTTGGCAGAGTGACAGTGAAGGGCTCTGTGTTTGTCTCGAAGAGAC 4631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2315 GGAGGCACTGACAGTTGGCAGAGTGACAGTGAAGGGCTCTGTGTTTGTCTCGAAGAGAC 2374
OY 4632 AAATACATTTTACAAGAGTCACTTCTGGGGTCTTGGTGTGACAGCCCAATTAAGCCT 4691
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2375 AAATACATTTTACAAGAGTCACTTCTGGGGTCTTGGTGTGACAGCCCAATTAAGCCT 2434
OY 4692 GGTGTATAGCTCGTGTTCAGGTTTGTACTTGTGATGAGGAATGATGAGAAATTAAT 4751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2435 GGTGTATAGCTCGTGTTCAGGTTTGTACTTGTGATGAGGAATGATGAGAAATTAAT 2494
OY 4752 TAAATGAGCGGAGAGCAGTGAAGCANTCAACTTAATGAATGAAAGTGGGTAAG 4811
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2495 TAAATGAGCGGAGAGCAGTGAAGCANTCAACTTAATGAATGAAAGTGGGTAAG 2554
OY 4812 ATTTAGCATGCTGCAAAATTAATGACAGCAATCAACGAAAGACATGTTCCACACTACAG 4871
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2555 ATTTAGCATGCTGCAAAATTAATGACAGCAATCAACGAAAGACATGTTCCACACTACAG 2614
OY 4872 CTAATGCAAACTTGCGATTTTGTGTATTTTGTGTATTAAGCTTTTAAGTCTGACTGAC 4931
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2615 CTAATGCAAACTTGCGCAATTTTGTGTATTTTGTGTATTTTGTGTATTTTGTGTATTT 2657
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
OY 4932 AAATCTGATTAAGTGTGATACGTATGACATTTGTTAAAAATTAACCTGTGACTTATT 4991
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2658 GGAATTCGTAGTATGAGTGTGACATGCTATGACATTTGTTAAAAATTAACCTGTGACTTAA 2717
OY 4992 TTGATTT 4998
    || ||
Db 2718 TTGATTT 2724
```

RESULT 8
AAQ12554
ID AAQ12554 standard; DNA; 2756 BP.

XX
AC AAQ12554;
XX
DT 23-SEP-1991 (first entry)
XX
DE Encodes Plasminogen mutelin T17 with thrombin cleavage site.
XX
KW protease; fibrinolysis; blood clotting; ss.

XX
FH Key Location/Qualifiers
FT mutation 179..1801
FT /tag= a
FT /note= "Ser codon (AGT) inserted"

XX
PN M09109118-A.

XX
PD 27-JUN-1991.

XX
PF 07-DEC-1990; 90WO-G001912.

XX
PR 07-DEC-1989; 89GB-0027722.

XX
PR 07-DEC-1990; 90WO-GB01911.

XX
PA (BRB1-) BRIT BIO-TECHN LTD.

XX
PI Dawson KM, Edwards RM, Forman JM;

XX
DR WPL: 1991-208145/28.

XX
DR P-PSDB; AAR12945.

XX
PT Activatable fibrinolytic and antithrombic proteins - activated by
e.g. factor Xa, thrombin or activated protein C

XX
PS Claim 16; Fig 2 and Fig 5; 73pp; English.

XX
CC This sequence encodes a plasminogen mutant cleavable by thrombin.
CC Activation is localised to the thrombus because cleavage to plasmin
CC is by an enzyme of the blood clotting pathway. Compositions
CC comprising the mutant plasminogen encoded by this sequence are used
CC for treatment or prevention of thrombosis, etc.
XX
CC See AAQ12542-Q12558.

XX
SQ Sequence 2756 BP; 766 A; 671 C; 696 G; 623 T; 0 other;

Query Match 24.4%; Score 1219.2; DB 12; Length 2756;
Best Local Similarity 80.7%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 278; Indels 83; Gaps 4;

```
OY 3159 TATGAGGCAATATCCACACACTGTACAGAGAGGACCTGTCAATTTGTGATCTATG 3218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 911 TATGCGGGAAATGAGTGTATACCGTGTCCGGGACACACTGTACACACTGAGTGAACAG 970
OY 3219 ATACCACTGCGATCAGAGAGACCCAGAAAACTTACCCAAATGCTGCGCTGACGAGAAC 3278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 971 ACCCTCAGACACATTAACAGAGACACCAAAACTTTCCCTGCAAAAATTTGATGAAAC 1030
OY 3279 TACTGAGGAATCCAGATTTCTGGAAACAACCTGGTGTATACCAACGATCCGTGTG 3338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1031 TACTGCCCAATCTGTACGAAAAAAGGCCCATGTGTGCTATCAACCAACGCAAGTG 1090
```


XX
PA (BRBI-) BRIT BIO-TECHN LTD.
XX
PI Dawson KM, Edwards RM, Forman JM;
XX
DR WPI; 1991-208145/28.
XX P-PSDB; AARI2939.
XX
PT Activatable fibrinolytic and antithrombotic proteins - activated by
PT e.g. factor Xa, thrombin or activated protein C
XX
PS Claim 16; Fig 2 and Fig 5; 73pp; English.
XX
CC This sequence encodes a plasminogen mutant cleavable by thrombin.
CC Activation is localised to the thrombus because cleavage to plasmin
CC is by an enzyme of the blood clotting pathway. Compositions
CC comprising the mutant plasminogen encoded by this sequence are used
CC for treatment or prevention of thrombosis, etc.
CC See AA012542-Q12558.
XX
SQ Sequence 2756 BP; 765 A; 671 C; 697 G; 623 T; 0 other;

Query Match 24.3%; Score 1216; DB 12; Length 2756;
Best Local Similarity 80.6%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 280; Indels 83; Gaps 4;

QY 3159 TATCGAGGATATCCCTCCACGCTGTCCACAGAGAACCTGTCAATCTTGTCATCTATG 3218
DB 911 TATCGGAGGATGTGGCTTCTTACCTGTCCGGCACACCTGTGCAGCTGAGTGACAG 970
QY 3219 ATACCACTGGATCGATCAGAGACCCAGAAAACTACCCAAATGCTGGCTGACCGAGAAC 3278
DB 971 ACCCTTCACACACATTAACGAGACACAGAAAACTTCCCTGCAAAAATTGGATGAAAAAC 1030
QY 3279 TACTGAGGAATTCATCTGTGGAAACACCCCTGTGTACACGAGTCCGTGTGTC 3338
DB 1031 TACTGCCGCAATCTTACGGAAGAAAGGCCCATGTGTCCATACACACAGCCAAAGTG 1090
QY 3339 AGGTGGAGATGCTGCAATCTGACACAAATGCTCAGAAACGAATGAGTGTCTAGAGACT 3398
DB 1091 CGGTGGAGATGCTGTAAGTATCCGTCCTG----- 1119
QY 3399 CCACGTGTGTCCATGTCACAGCANTGAGAGGCTCATTTGTGAAGACCACTGAGCAA 3458
DB 1120 -----TGACTCCCTCCCACTATCCAGGACAAATTTGGCTCCACAGCACCTGAGACTA 1174
QY 3459 ACCCTGTGTGTCGGCAGCTGCTACCATGTATGGCCAGATTACAGGACATTTCTCC 3518
DB 1175 ACCCTGTGTGTCGGCAGGACTGCTACCATGTGTGATGACAGAGCTACCGAGGCATCTCC 1234
QY 3519 ACCACTGTACAGGAGGATGCAATCTGTGTCAATGCATGACACACACACCCGACATCAG 3578
DB 1235 ACCACACACACAGGAGAAAGTGTGATGTGTATGTATGACACACACACCCGACACAG 1294
QY 3579 AGGACCCAGAAAACTACCAATGATGCGCTGACAAATGAATCTACTGAGGAATCCAGAT 3638
DB 1295 AAGACCCAGAAAACTACCAATGATGCGCTGACAAATGAATCTACTGAGGAATCCAGAT 1354
QY 3639 GCGGATACAGGCGCTGTGTGTGTACGATGAGCCACACATCAGGTGGAGATTAATGCAAC 3698
DB 1355 GCGGATACAGGCGCTGTGTGTGTGTACGATGAGCCACACATCAGGTGGAGATTAATGCAAC 1414
QY 3699 CTGACGCGATGCTCAGACAGAGAGGACTGTGTGTCTCCGATGCTCATCAGAT 3758
DB 1415 CTGAAAAAATGCTCAGGACAGAGAGATGTGTGACAGCTCCGCGCTGTCTCTCTT 1474
QY 3759 CCAAGCCTAGGCGCTCTCTCTGACAGAGCTGTATGTTGGGATGGGAAGATACCGG 3818
DB 1475 CCGATGTGTAGAGCTCTCTCTCGAAGAGCTGTATGTTGGGATGGGAAGATACCGG 1534
QY 3819 GGCAGAGAGCAACCACTGTACTGTGAGCGCCATGCCAGATGGGCTGCCAGAGACCC 3878
DB 1535 GGCAGAGGCGCACACTGTACTGTGAGCGCCATGCCAGAGACTGGGCTGCCAGAGACCC 1594

QY 3879 CATAGACAGACAGCTTCATTCACAGGAGACAATTAATGGGCGAGTCTGGAATAAATTTAC 3938
DB 1595 CATAGACAGACAGCTTCATTCACAGGAGACAATTCAGGCGGCGGTCTGAAAAAATTTAC 1654
QY 3939 TGCCGTAAACCTGATGTGTGACATCAATGATCCCTGGTGTCTACCAATGAATCCAGAAAA 3998
DB 1655 TGCCGTAAACCTGATGTGTGATGTAGTGTGCTCCCTGGTGTCTACCACAATATCCAGAAAA 1714
QY 3999 CTTTGTACATGCTGATATATCCCTCTCTGTGCATCCTCTCTCAATTTGATGTGGGAAGCT 4058
DB 1715 CTTTGTACATGCTGATGT 1774
QY 4059 CAGGTGAGCCGGAAGAATGCTCTG--AAGCATTTGTAGGGGGGTGTGTGGCCACCA 4115
DB 1775 CAGGTGAGCCGGAAGAATGCTGT 1834
QY 4116 CATTCCTGGCCCGGCAAGTCACTCAGAACAGGTTTGGAAAGCACTCTGTGGAGGC 4175
DB 1835 CATTCCTGGCCCGGCAAGTCACTCAGAACAGGTTTGGAAAGCACTCTGTGGAGGC 1894
QY 4176 ACCTTAATATCCCGAGAGTGGTGTGCTGCTGCTCACTGCTTGAAGAGTCTCAAG 4235
DB 1895 ACCTTAATATCCCGAGAGTGGTGTGCTGCTGCTGCTTGAAGAGTCTCAAG 1954
QY 4236 CCTTCATCCTCAAGAGTCACTCCTGGGCGACACCAAGAGTCACTGAGTCAATGTT 4295
DB 1955 CCTTCATCCTCAAGAGTCACTCCTGGGCGACACCAAGAGTCACTGAGTCAATGTT 2014
QY 4296 CAGGAAATAGAAAGTGTGTAGGCTGTCTTGTGAGCCCAACAAGAGATATTGCTTGCTA 4355
DB 2015 CAGGAAATAGAAAGTGTGTAGGCTGTCTTGTGAGCCCAACAAGAGATATTGCTTGCTA 2074
QY 4356 AAGCTAAGCAAGGCTGCGGCTCATCAGCAAGTAATCCAGCTTGTCTGCATCCCA 4415
DB 2075 AAGCTAAGCAAGTCTGCGGCTCATCAGCAAGTAATCCAGCTTGTCTGCATCCCA 2134
QY 4416 GACTACATGCTACCGCAGCACTGAATGTTCATCATCTGCTGGGGAAGAAACCAAGT 4475
DB 2135 AATTATGAGTGTGCTGACCGCAGCAATGTTCATCATCTGCTGGGGAAGAAACCAAGT 2194
QY 4476 ACCTTGGAGCTGCGCTTTCAGAGAAAGCCAGCTCTTGAATGAGAAATGAAGTGTG 4535
DB 2195 ACCTTGGAGCTGCGCTTTCAGAGAAAGCCAGCTCTTGAATGAGAAATGAAGTGTG 2254
QY 4536 AATCATATAAGT-----AATTTGTCTGAGCATTTG 4568
DB 2255 AATCATATAAGT-----AATTTGTCTGAGCATTTG 2314
QY 4569 GCCAGAGGCACTGACATTTGCCAGGGTGAAGAGTGAAGTCAATCCAGCACTGTCTGGGCAATTTG 2374
DB 2315 GCCAGAGGCACTGACATTTGCCAGGGTGAAGAGTGAAGTCAATCCAGCACTGTCTGGGCAAG 2434
QY 4639 GACAAATACATTTTACAAGAGTCACTCTTGGGGTCTTGGCTGTGCAAGCCCAATAG 4688
DB 2375 GACAAATACATTTTACAAGAGTCACTCTTGGGGTCTTGGCTGTGCAAGCCCAATAG 2434
QY 4689 CCGGTGTCTATGCTGTGCTGTTCAGAGTGTGTACCTGTGATGAGGAGATGAGGAAT 4748
DB 2435 CCGGTGTCTATGCTGTGCTGTTCAGAGTGTGTACCTGTGATGAGGAGATGAGGAAT 2494
QY 4749 AATTAAATGAGGAGAGAGAGTGAAGCATCACTTAATGAAGCTGAACGTGGTA 4808
DB 2495 AATTAAATGAGGAGAGAGAGTGAAGCATCACTTAATGAAGCTGAACGTGGTA 2554
QY 4809 AGGATTAATGCTGTGGAATTAATGAAGCATCAAGAGATGAAGAGACAGTGTCCAGCTAC 4868
DB 2555 AGGATTAATGCTGTGGAATTAATGAAGCATCAAGAGATGAAGAGACAGTGTCCAGCTAC 2614
QY 4869 CAGCTATGCCAAACCTTGGCATTTTGTATTTTGTGTATTAAGCTTTTAAGGTCTGACT 4928
DB 2615 CAGCTATGCCAAACCTTGGCATTTTGTGTATTTTGTGTATTTTGTGTATTTTGTGTACT 2657

Db	1595	CATGACACACACATTTTCTACTCCAGAGACAAATCCAGGGGGGCTGTGAAAAAAATTAC	1654
QY	3939	TGCCGTAAACCTGATGTGTGATCATCAATATGGTCCCTGGTGTGTACACATGAATGAATCAAGAAA	3988
Db	1655	TGCCGTAAACCTGTATGTGTATGTATAGGTGGTCCCTGGTGTGTACACGAATATCCAGMAAA	1714
QY	3999	CTTTTGTACATCTGTATATATCCCTCTCTGTGTGATCCTCTTCAATTTGATTTGGGAAGCTT	4058
Db	1715	CTTTACAGACACTGTGTATGTATGTCCCTCAGTGTGGCGGCCCTTCTCAATTTGATTTGGGAAGCTT	1774
QY	4059	CAAGTGGAGCCGAAGAATATCTCTGGA---AGCATTTGTAGGGGGGGTGTGGCCCAACCA	4115
Db	1775	CAAGTGGAGCCGAAGAATATGTATTGTATACCTAGAGTTGTGGGGGGGTGTGTGGCCCAACCA	1834
QY	4116	CATTCTGGCCCTGGCAAGTCACTCTCAGAACAGGTTTTGAAAAGCACTTGTGTGAAGC	4175
Db	1835	CATTCTGGCCCTGGCAAGTCACTCTTAGAACAGGTTTGGAAATGCACCTCTGTGTGAAGC	1894
QY	4176	ACCTTAATATCCCGAGAGTGGGGTGTGACGTGGCGTCACTGCTGTGAAGAAGTCCCAAG	4235
Db	1895	ACCTTGTATATCCCGAGAGTGGGGTGTGTACTGTCTGGCCACCTGTGTGAAGAAGTCCCAAG	1954
QY	4236	CCCTTCATCTCAACAAGTCACTCTGTGGTGCACACCAAGAAGTGAACCTCGAATCTCATTT	4295
Db	1955	CCCTTCATCTCAACAAGTCACTCTGTGGTGCACACCAAGAAGTGAATCTCGAACCCAGTGT	2014
QY	4296	CAGAAATTAAGAAGTGTATAGGCGTTCCTGGAGCCCAACAGACAGATATTCCTTGCTA	4355
Db	2015	CAGAAATTAAGAAGTGTATAGGCGTTCCTGGAGCCCAACAGAAATATATTCCTTGCTA	2074
QY	4356	AAGTAAAGCAGGCGCTGCCGTCACTGTGACAAAGTAATGCCAGCTTGTGTGCCATCCCA	4415
Db	2075	AAGCTAAGCAGTCTGCCGTCACTGTGACAAAGTAATCCAGCTTGTGTGCCATCCCA	2134
QY	4416	GACTAATAGGTACACCGCAGAGACTGAAATGTATATACAGCGCTGGGGGAAAAACCAAGT	4475
Db	2135	AATTATGTGTCTGTGACCGGACCGAATGTATATACAGCGCTGGGGGAAAAACCAAGT	2194
QY	4476	ACCTTTGGGAGCTGGCTTCTCAAGGAAGCCAGCTCCTTGTATTATGAGATGAATGTGC	4535
Db	2195	ACCTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCTGTGATGTGAATAAATGTGTC	2254
QY	4536	AATCACTATATAGT-----AATTTGTGCTAGCATTTG	4586
Db	2255	AATGCTATGATGTTTCTGAATGGAAGATCCAAATCCACCGAACTCTGTCTGGGATTTG	2314
QY	4589	GCCAGAGGCACTGCAGTGTCCAGAGGTGACAGTGAAGGAGCCCTGTGGTTGCTTCAGAG	4628
Db	2315	GCCGAGGCACTGCAGTGTCCAGAGGTGACAGTGAAGGAGTCTCTGTGTTGCTTCAGAG	2374
QY	4629	GACAAATATCATTTTACAGAGATCACTCTTGGGGTCTTGGCTGTGCACGCCCAATTAAG	4688
Db	2375	GACAAATATCATTTTACAGAGATCACTCTTGGGGTCTTGGCTGTGCACGCCCAATTAAG	2434
QY	4689	CCGTGTGTCAATGCTGTCTGTTCCTTCAAGTTTGTATCTTGGATTGAAGGAATGATGTAAT	4748
Db	2435	CCGTGTGTCTATGTCTGTGTTCCTTCAAGTTTGTACTTGGATTGAAGGAATGATGTAAT	2494
QY	4749	AATTAAATTTGACGGGAGACAGAGTGAAGCATCAACCTACTTGAAGAGCTGMAACGTGGGTA	4808
Db	2495	AATTAAATTTGACGGGAGACAGAGTGAAGCATCAACCTACTTGAAGAGCTGMAACGTGGGTA	2554
QY	4809	AGGATTTAGCATGTGTGGAATTAATTAACAGCAATCAAAAGAGACACATGTTCACAGTAC	4868
Db	2555	GGGATTTAGCATGTGTGGAATTAATCTGGCAATATCAAAAGAGACACATGTTCACAGTAC	2614
QY	4869	CAGATATGCCAAACCTTGGCATTTTGGTATTTTGTGTATAGCTTTTAAAGTGTGACT	4928
Db	2615	CAGCTAGCGGAAACCTGGCATTTTGTGTATTT-----TTCGTGACT	2657
QY	4929	GACAAATTTCTGTATTAAAGTGTATAGCTATGACATTTGTAAAAATAAATCTGCACCTT	4988

Dd	2658	GCTGATTCCTGTAGTAAAGCTGCATAGCTATTGTAAAAATAACTCTGTACTT	2717
Oy	4989	ATTTTTGATTT	4998
Dd	2718	AACTTTGATT	2727
 RESULT 12			
ID	AAQ12546		
NC	AAQ12546	standard; DNA; 2762 BP.	
XX	AAQ12546;		
XX	AAQ12546;		
DT	23-SEP-1991	(first entry)	
DE	Encodes Plasminogen mutetin X6 with factor Xa cleavage site.		
XX			
XX	protease; fibrinolysis; blood clotting; ss.		
FH	Key	location/Qualifiers	
FT	mutation	1796..1807	
FT		/tag= a	
FT		/note= "replaces CCT (Pro codon) with four codons"	
FT	mutation	1814..1816	
FT		/tag= b	
FT		/note= "replaces GTT (Val codon) with ATT (Ile)"	
FT	CDS	65..2509	
FT		//tag= c	
FT		/product= modified plasminogen	
XX			
PN	W09109118-A.		
PD	27-JUN-1991.		
XX			
PE	07-DEC-1990;	90WO-G001912.	
XX			
PR	07-DEC-1989;	89GB-0027722.	
PR	07-DEC-1990;	90WB-GB01911.	
PA	(BRRI-) BRIT BIO-TECHN LTD.		
XX			
PI	Dawson KM, Edwards RM, Forman JM;		
DR	WPI: 1991-208145/28.		
XX	P-PSDB; AARI2937.		
PT	Activatable fibrinolytic and antithrombic proteins - activated by		
XX	e.g. factor Xa, thrombin or activated protein C		
PS	Claim 16; Fig 2 and Fig 4; 73pp; English.		
CC	This sequence encodes a plasminogen mutant cleavable by Factor Xa.		
CC	Activation is localised to the thrombus because cleavage to plasmin		
CC	is by an enzyme of the blood clotting pathway. Compositions		
CC	comprising the mutant plasminogen encoded by this sequence are used		
CC	for treatment or prevention of thrombosis, etc.		
CC	See AAQ12542-Q12558.		
SQ	Sequence 2762 BP; 770 A; 671 C; 697 G; 624 T; 0 other:		
 Query Match 24.2%; Score 1211.6; DB 12; Length 2762; Best Local Similarity 80.4%; Pred. No. 0; Matches 1508; Conservative 0; Mismatches 279; Indels 89; Gaps 4;			
Oy	3159	TATCGAGCATAATCTCTCCACACATGTCTACAGAGAAGACCTGTCAATCTGTGTCATATG	3218
Dd	911	TATCGCGGAATGTGGCTGTATCCGTGTGCCGGCACACATGTCTGACACTGGAGTCACAG	970
Oy	3219	ATAACACACATGGCANCAAGACCACAGAAAAGTAACCAATATGCTGGCTGACCGAGAAC	3278
Dd	971	ACCCTCATACACATAAACAGACACACAGAAAATCTTCCCTGCACAAAATTGGATGAANAAC	1030
Oy	3279	TACTGACAGAAATCCAGATTCCTGGGAAACACACCCTGGTGTATTACACACCGATCCGTGTG	3338


```

Db 2075 TTGCTAAAGCTAAGCAGTCCGCGTCATCTACACAAAGTAATCCAGCTTGCTGCA 2134
OY 4410 TCCCAGACTACATGCTACGCGGAGAGTAAATGTAACATCACTGGCTGGGGAAC 4469
Db 2135 TCCCAAAATTTATGTGCTGCGACCGAGCGAATGTTTACTGCTGGGGAAC 2194
OY 4470 CAAGGTACCTTTGGGACTGGCTTCTCAAGGAAGCCAGCTCTGTTATGAGAATGAA 4529
Db 2195 CAAGTACTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGATGATGAATAAA 2254
OY 4530 GTGTCATCTACTATAGT-----ATATTGTGCTGAG 4562
Db 2255 GTGTGCATCTGCTATGAGTTTCTGAATGGAAGAGTCCATCCACCGAAGCTGTGCTGGG 2314
OY 4563 CATTTGGCCAGAGCAGTACAGTATGCGAGGAGTACAGTGGAGGGCCCTGGTTGCTTC 4622
Db 2315 CATTTGGCCGAGGACACTGACAGTGGCCAGAGGTACAGTGGAGGCTCTGTGTTGCTTC 2374
OY 4623 GAGAAGGACAAATACATTTTACAAGAGTCACTTTGGGGTCTTGGCTGTGCAGCCCC 4682
Db 2375 GAGAAGGACAAATACATTTTACAAGAGTCACTTTGGGGTCTTGGCTGTGCAGCCCC 2434
OY 4683 AATTAAGCCTGCTCTATGCTCGTGTTCAGAGTTTGTACTTGTGATTGAGGAGATGATG 4742
Db 2435 AATTAAGCCTGCTCTATGCTCGTGTTCAGAGTTTGTACTTGTGATTGAGGAGATGATG 2494
OY 4743 AGAATTAATTTATGAGCGGAGAGAGTGAAGTCAACCTTACTTAAGAGTGAAG 4802
Db 2495 AGAATTAATTTATGAGCGGAGAGAGTGAAGTCAACCTTACTTAAGAGTGAAG 2554
OY 4803 TGGGTAAGATTATGATGCTGTAATTAATAGACAGCAATCAAGAGACACTGTTCC 4862
Db 2555 TGGGTAAGATTATGATGCTGTAATTAATAGACAGCAATCAAGAGACACTGTTCC 2614
OY 4863 AGCTACAGCATATGCCAAACCTTGGCTTTTGTATTTTGTGTAACTTTTAAGT 4922
Db 2615 AGCTACAGCATATGCCAAACCTTGGCTTTTGTGTAACTTTTAACTTTTAACTCT 2657
OY 4923 CTGACTAGCAAAATCTGATATTAAGTGTCAATAGCTATGACATTTGTTAAATTAACCT 4982
Db 2658 CTGACTAGCTGATCTGTAGTAAGGTACATAGCTATGACATTTGTTAAATTAACCTCT 2717
OY 4983 GCACCTATTTTGAATTT 4998
Db 2718 GTACTTAACCTTTGATTT 2733

RESULT 15
AAQ12545
ID AAQ12545 standard; DNA: 2762 BP.
AC AAQ12545;
XX
XX 23-SEP-1991 (first entry)
DE Encodes Plasminogen mutetin X5 with factor Xa cleavage site.
KM protease; fibrinolysis; blood clotting; ss.
XX
XX Key Location/Qualifiers
FH 1796..1807
FT mutation
FT /tag= a
FT /note= "Replaces CCT (Pro codon) with four codons"
FT CDS 65..2509
FT /tag= b
FT /product= modified plasminogen
XX
XX WO9109118-A.
XX 27-JUN-1991.
XX
XX 07-DEC-1990; 90MO-G001912.
```

```

XX
PR 07-DEC-1989; 89GB-0027722.
PR 07-DEC-1990; 90MO-GB01911.
XX
XX (BRBI-) BRIT BIO-TECHN LTD.
XX
XX Dawson KM, Edwards RM, Forman JM;
XX
XX MPI: 1991-208145/28.
XX P-PSDB; AAR12936.
XX
XX Activatable fibrinolytic and antithrombic proteins - activated by
XX e.g. factor Xa, thrombin or activated protein C
XX
XX Claim 16; Fig 2 and Fig 4; 73pp; English.
XX
XX This sequence encodes a plasminogen mutant cleavable by Factor Xa.
XX Activation is localised to the thrombus because cleavage to plasmin
XX is by an enzyme of the blood clotting pathway. Compositions
XX comprising the mutant plasminogen encoded by this sequence are used
XX for treatment or prevention of thrombosis, etc.
XX See AAQ12542-Q12558.
XX
SQ Sequence 2762 BP; 769 A; 671 C; 698 G; 624 T; 0 other;
XX
Query Match 24.2%; Score 1210; DB 12; Length 2762;
Best Local Similarity 80.3%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 280; Indels 89; Gaps 4;
OY 3159 TATCGAGCATATCTCCACCACTGTACAGAGGACCTGCAATCTTGATCTATG 3218
Db 911 TATCGGCGGAATGTGCTGTATACCTGTCGCGGACACCTGTACAGCATGATGACAG 970
OY 3219 ATTCACACTGGATACAGAGACCCAGAAATCACTACCCAAATGCTGGCTGACCGAAG 3278
Db 971 ACCCTTCACACATATACAGAGACACCAAGAACTTCTCTGCAAAATTTGGATGATAAAC 1030
OY 3279 TACTGAGGAATCCAGATTTCTGGGAAACAACCTGTGTTTACACACGATCCGTGTG 3338
Db 1031 TACTGCGGCAATCTTACGGAAGAAAGGCCCATGTGTGCATACACCAAGCCCAAGTG 1090
OY 3339 AGGTGGAGTACTGCATCTGACACATGCTCAAGAAACGAAATCAGTGTCTTAGAGACT 3398
Db 1091 CGGTGGAGTACTGTAATACCGTCTG----- 1119
OY 3399 CCCACTGTGTCCCACTTCCAGATGAGAGCTCATTTGAGACGACACCACTAGCAA 3458
Db 1120 -----TGACTCCTCCCACTATCCACGGAACAATTTGGCTCCACAGACACCACTAGCTA 1174
OY 3459 ACCCTGTGTGTCGCGAGTGTACCATGATGTAATGGCCAGAGTTATCAGAGCATTTCTCC 3518
Db 1175 ACCCTGTGTGTCGAGAGTGTACCATGATGTAATGGCCAGAGTGTACGAGCATCTCTCC 1234
OY 3519 ACCACTGTACAGAGAGACATGCAATCTTGTGATCCATCCAGACACACCGGCATGAG 3578
Db 1235 ACCACACCAAGAAAGTGTGATGCTGTATGATCATACACACACCGGCATGAG 1294
OY 3579 AGGACCCCAAGAAATCAACCAATGATGCTGGAATGAATTAATGAGAAATCCAAAT 3638
Db 1295 AAGACCCCAAGAAATCAACCAATGATGCTGGAATGAATTAATGAGAAATCCAAAT 1354
OY 3639 GCCGATACAGGCCCTTGTGTTTACATGAGACCCACAGATGAGTGGAGTACTCTCAAC 3698
Db 1355 GCCGATTAAGGCCCTTGTGTTTACATGAGACCCACAGATGAGTGGAGTACTCTCAAC 1414
OY 3699 CTGACGCGATGCTCAGACACAGAGGAGCTGTGCTCTCTCCGACGTGCATCCAGGTT 3758
Db 1415 CTGAAAAAATGCTCAGGAAGAGAGAGCTGTGAGACCTCCGCTGTGTCTGCTT 1474
OY 3759 CCAAGCCTAGGGCTCTCTTAACAAGACTATGTTTGGGAATGGGAAGATATCCGG 3818
Db 1475 CCAAGTGTAGAGACTCTTCCAGAAAGACTGTATGTTGGGAATGGGAAGATATCCGA 1534
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 06:19:18 ; Search time 148 Seconds
(without alignments)
10360.701 Million cell updates/sec

Title: US-09-923-515-3_COPY_8939_13938

Perfect score: 5000

Sequence: 1 ctccgactgtaccgccggtt.....ctgcattattgtattga 5000

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1237	24.7	2753	1	US-07-854-603-1
2	1236.8	24.7	2679	6	5200340-7
3	1110.6	22.2	2296	1	US-07-750-080A-18
4	1110.6	22.2	2296	4	US-08-651-472-18
5	1110.6	22.2	2296	4	US-08-358-928-18
6	1104.8	22.1	2497	1	US-08-643-219-12
7	1104.8	22.1	2497	2	US-09-131-995-12
8	1104.8	22.1	2497	2	US-08-832-087B-12
9	1104.8	22.1	2497	3	US-08-851-350-12
10	1104.8	22.1	2497	4	US-09-132-154-12
11	697.6	14.0	1724	6	5200340-5
12	557.8	11.2	690	1	US-08-379-621-1
13	557.8	11.2	690	2	US-08-889-078-1
14	257.2	5.1	1134	4	US-09-206-059-29
15	90.8	1.8	2172	4	US-08-030-410-2
16	89.2	1.8	2184	1	US-07-815-333A-1
17	89.2	1.8	2288	1	US-08-290-937B-4
18	89.2	1.8	2288	1	US-07-838-410-2
19	85.4	1.7	2219	1	US-07-882-925A-1
20	85.4	1.7	2219	1	US-07-882-925A-2
21	85.4	1.7	2219	1	US-08-184-012C-1
22	85.4	1.7	2219	1	US-08-184-012C-2
23	84.2	1.7	2216	2	US-08-666-082B-2
24	84.2	1.7	2232	1	US-08-334-177-1
25	84.2	1.7	2232	5	PCT-US95-13830-1
26	84.2	1.7	2262	1	US-07-882-925A-7
27	84.2	1.7	2262	1	US-08-184-012C-7

28	80.4	1.6	645	3	US-08-985-526-22	Sequence 22, Appl
29	80.4	1.6	1284	3	US-08-985-526-24	Sequence 24, Appl
30	80	1.6	2188	1	US-07-882-925A-4	Sequence 4, Appl
31	80	1.6	2188	1	US-08-184-012C-4	Sequence 4, Appl
32	68	1.4	2021	1	US-07-882-925A-3	Sequence 3, Appl
33	68	1.4	2021	1	US-08-184-012C-3	Sequence 3, Appl
34	64.6	1.3	2544	4	US-09-518-046-3	Sequence 3, Appl
35	63	1.3	2413	4	US-09-518-046-1	Sequence 1, Appl
36	62.2	1.2	1130	4	US-09-386-653A-8	Sequence 8, Appl
37	62	1.2	654	4	US-09-280-116-57	Sequence 57, Appl
38	62	1.2	1110	4	US-09-386-653A-1	Sequence 1, Appl
39	60.8	1.2	6100	1	US-07-882-925A-6	Sequence 6, Appl
40	60.8	1.2	6100	1	US-08-184-012C-6	Sequence 6, Appl
41	60.6	1.2	2038	4	US-09-008-271A-18	Sequence 18, Appl
42	60.6	1.2	2079	4	US-09-656-002-1	Sequence 1, Appl
43	60.4	1.2	1109	4	US-09-088-651-1	Sequence 1, Appl
44	60.2	1.2	2416	4	US-09-261-416-1	Sequence 1, Appl
45	59.4	1.2	3398	5	PCT-US95-08493-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
: Sequence 1, Application US/07854603
: Patent No. 5637492
: GENERAL INFORMATION:
: APPLICANT: Dawson, Keith M
: APPLICANT: Edwards, Richard M
: TITLE OF INVENTION: Forman, Joan M
: TITLE OF INVENTION: Activatable fibrinolytic and
: NUMBER OF SEQUENCES: 40
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dr. John J. McDonnell
: STREET: Ten South Wacker Drive, Suite 3000
: CITY: Chicago
: STATE: IL
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/854, 603
: FILING DATE: 19901207
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: McDonnell, John J
: REGISTRATION NUMBER: 26,949
: REFERENCE/DOCKET NUMBER: 92,338
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-715-1000
: TELEFAX: 312-715-1234
: TELEX: 910-221-5317
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2753 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: 1..2753
: OTHER INFORMATION: /note="Fig. 2 Plasminogen CDNA
: OTHER INFORMATION: sequence"
: FEATURE:
: NAME/KEY: sig.peptide
: LOCATION: 65..121

[illegible][illegible]

QY	4539	CACATTAAGT-----	-----ATATTGGTCGAGCATTTGGCC	4571
Db	1980	CGCTATGAGTTTCTGATGGAAGAGTCCAAATCCACCGAACTCTGTGGGGATTTTGGCC		2039
QY	4572	AGAGGCACTGCACAGTGTCCAGGAGGTGACAGTGGAGGGCCCTGTGGTTGCTTCAGAGAAGAC		4631
Db	2040	GGAGGCACTGCACAGTGTCCAGGAGGTGACAGTGGAGGGTCCCTGTGGTTGCTTCAGAGAAGAC		2099
QY	4632	AAATTCATTTTACAAAGAGACTCACTTCTTGGGGTCTTGGCTGTGCACGCCCAATTAAGCCT		4691
Db	2100	AAATTCATTTTACAAAGAGAGTCACTTCTTGGGGTCTTGGCTGTGCACGCCCAATTAAGCCT		2159
QY	4692	GGTGCCTATNGCTCGAGTTTCAAGGTTTGTATCTTGATTTGAGAGGATGATGAAGAATAT		4751
Db	2160	GGTGCCTATNGCTCGAGTTTCAAGGTTTGTATCTTGATTTGAGAGGATGATGAAGAATAT		2219
QY	4752	TAAATTGACGGGAGACAGAGTGAACGA	4778	
Db	2220	TAAATTGACGGGAGACAGAGTGAACGA	2246	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

US-08-651-472-18
; CLONE: PN29pt-LPg
; IMMEDIATE SOURCE:

Query Match	22.2%;	Score 1110.6;	DB 3;	Length 2296;
Best Local Similarity	81.4%;	Pred. No. 0;		
Matches 1340;	Conservative 0;	Mismatches 244;	Indels 63;	Gaps 2

QY	3159	TATGAGGCAATATCTCCACCACTGTGCACAGGAAGAACTGTGCATCTTTGGTCACTCTATG	3218
Db	636	TATCCCGGGAAATGTGGCTGTATTACCGTTTCCGGGCACACCTGTGCACACTGCTAGCGTGGTGCACAG	695
QY	3219	ATACCAACACTGGCAATCTCAGAGAGACCCAGAAATACTACCCAAATGTGGCTGTGACCGAGAAC	3278
Db	696	ACCCTTCACACACATTAACAGGACACCGAATACTTCCCTGCAGAAAATTTTGGATGAAAC	755
QY	3279	TACTGCAGAAATCCAGATTTCTGGGAAACAACCTGGTGTGTACCAACCGATCCGTGTG	3338
Db	756	TACTTCGCCCAATCTCTGAAGGAAAAAGGGCCCATGGTGGCAATCAACCAAGCCAACTG	815
QY	3339	AGGTGGAGTACTCCTCAATCTGACACAATGCTCAGAAAACGAATCAGTGTCTTAGACT	3398
Db	816	CGGTGGGAGTACTGTAAAGATACGCTCTG-----	844
QY	3399	CCCACTGTATTCCAGTTCCAGAGATGAGGCTCATTTGGAAGCAGCACCACTGAGAA	3458
Db	845	-----TGACTCTCTCCCACTGTCCACGGAAATTTGGCTCCACAGCACCACTTAGCTA	899
QY	3459	ACCCTGTGTGTCGGGACAGTGTACCATGTAAATGGCAGAGATTAGAGCAATTTCC	3518
Db	900	ACCCTGTGTGTCAGAGACTGTACACAGGTATGAGACAGAGCTACCGAGCACTCTCC	959
QY	3519	ACCACTGTCAAGGAAGGACATGTCAATCTTGGTCACTCATGACACACACCGCATAG	3578
Db	960	ACCACCCACAGGAAGAAAGTGTCACTGTGTGTCACTATGTACACACACACCGGCACAG	1019
QY	3579	AGGACCCCAAGAAATACTCCAAATGATGGCTCATACTGAATCACTCAGAAATCCAGAT	3638
Db	1020	AAGACCCCAAGAAATCTACCAATAGCTGGCTCATACTGAATCACTCAGAAATCCAGAT	1079
QY	3639	GCCGATACAGGCGCTTGTTGTTTACCATGACCCACGACATCAGTGGGATCTGCAC	3698
Db	1080	GCCGATTAAGGCGCTCTGGTGTTTTACACAGACCCACGAGCTCAGGTGGGATCTGCAC	1139
QY	3699	CTGACGCGATCTCAGACACACAGGAAGGAGTGTGTGGCTCTCCCACTGTCAATCCAGAT	3758
Db	1140	CTGAAAAAATGTCTCAGAACGAAGGAGGTGTGTAGCACTCTCCGCTGTGTCTGTCTT	1199
QY	3759	CCAACTTAAGGGCCCTTCTGAAACAAGACTGTATGTTGGAAATGGAAAGATATCCGG	3818
Db	1200	CCAAATGTATGAGACTCTCTCCGAAGAAAGCTGTATGTTTGGAAATGGAAAGATATCCGA	1259
QY	3819	GGCAAGAAAGCAACACTGTTTACTGGAGCCATGCCAGGAATGGCTGCCAGAGACCC	3878
Db	1260	GGCAAGAGGGGACACTGTTACTGGAGCCATGCCAGGAATGGCTGCCAGAGACCC	1319
QY	3879	CATTACACAGACGCTTCATTCGAGGGAACAATAATGGGCAAGTGTGGAAAAAATTAC	3938
Db	1320	CATTACACAGATTTTCTACTCCAGAGACAAATTCACGGGGGCTGTGGAAAAAATTAC	1379
QY	3939	TGCCGTAAACCTGATGTGATCATTAATGTGCTCGCTGTGCTACACATGATTAACAAGAAA	3998
Db	1380	TGCCGTAAACCTGATGTGATGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1439
QY	3999	CTTTTGTACTGTGATATCCCTCTCTGTGTGCATCCTCTTCAATTTGATTTGTGGAAAGCT	4058
Db	1440	CTTTACGACTACTGTGATGTCTCTCAAGTGTGTGCGGCCCTTCAATTTGATTTGTGGAAAGCT	1499
QY	4059	CAAGTGGAGCGCAAAATGTCTGGAAGAAATGTATGGGGGGGTGTGGCCCAACCCACAT	4118
Db	1500	CAAGTGGAGCGCAAAATGTCTGTGGAAAGGTTGTGTGGGGGGGTGTGTGGCCCAACCCACAT	1558
QY	4119	TCTGTGGCCCTGGCAAGTCAAGTCTAGAAACAAGTTTGGAAAGCATCTTGTGTGAGGACAC	4178


```
|||||
Db 2360 AATACATTTTACAAGAGTACTTCTGGGGTCTTGCTGTCACGCCCAATTAAGCT 2419
QY 4692 GGTGCTATGCTCGTGTTCAGAGTTGTACTTGATTTGAGGAGATGATGAATAAT 4751
Db 2420 GGTGCTATGCTCGTGTTCAGAGTTGTACTTGATTTGAGGAGATGATGAATAAT 2479
QY 4752 TAATTGACGGAGACAG 4769
Db 2480 TAATTGACGGAGACAG 2497

RESULT 11
5200340-5
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MUIVHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINDEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO:5:
; LENGTH: 1724

Query Match 14.0%; Score 697.6; DB 6; Length 1724;
Best Local Similarity 80.2%; Pred. No. 2.2e+208;
Matches 858; Conservative 0; Mismatches 179; Indels 33; Gaps 2;
```

```
Db 1255 ACTGACAAAGTAATCCAGCTTGCTGCATCCCAATTAATGTGCTGCACGGACC 1314
QY 4440 GAATGTAAACATCACTGGCTGGGAGAAACCAAGTACCTTTGGAGCTGCCTTCAAG 4499
Db 1315 GAATGTTTCATCACTGGCTGGGAGAAACCAAGTACCTTTGGAGCTGCCTTCAAG 1374
QY 4500 GAAAGCCAGCTCTGTTGTAATGAGAAATGAAGTGTCAATCAATAAGT----- 4548
Db 1375 GAAGCCAGCTCCCTGTGATTAAGAAATAAAGTGTCAATCGATAGTTTCTGAATGA 1434
QY 4549 -----ATATTGCTGCTGAGCATTTGGCCAGAGGCACTGACAGTGGCAG 4592
Db 1435 AGAGTCAATCCACCAACTCTGTGCTGGCATTTGGCCAGGCACTGACATTTGGCAG 1494
QY 4593 GGTGACAGTGAAGGCGCTGTTGCTGAGAGGACAAATCATTTTACAAGAGTGC 4652
Db 1495 GGTGACAGTGAAGTCTCTGTGTTGCTTGGAGAGGACAAATCATTTTACAAGAGTGC 1554
QY 4653 ACTCTTGGGGCTTTGGCTGTGCAAGCCCAATTAAGCTGTGTATGCTGCTGTTCA 4712
Db 1555 ACTCTTGGGGCTTTGGCTGTGCAAGCCCAATTAAGCTGTGTATGCTGCTGTTCA 1614
QY 4713 AGCTTTGCTACTTGATTTGAGGAAATGATGAGAAATTAATTATTTGACGGAGACAGT 4772
Db 1615 AGCTTTGCTACTTGATTTGAGGAGATGATGAGAAATTAATTATTTGACGGAGACAGT 1674
QY 4773 GAAGATCAACCTACTTAAGCTGAAGCTGAAGCGTGGTAAAGATTATTCATGTC 4822
Db 1675 GACGCACCTGACTCACCTTAGAGCTGGAACGAGGTTAGGATTTAGCATGC 1724

RESULT 12
US-08-379-621-1
; Sequence 1, Application US/08379621
; Patent No. 5645833
; GENERAL INFORMATION:
; APPLICANT: DAWSON, Keith
; APPLICANT: GILBERT, Richard James
; TITLE OF INVENTION: INHIBITOR RESISTANT SERINE PROTEASES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: HALE AND DORR
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,621
; FILING DATE: 03-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/01632
; FILING DATE: 03-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Hollie L.
; REGISTRATION NUMBER: 31,321
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 942-8400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 690 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
```

```

? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..690
? OTHER INFORMATION:
? OTHER INFORMATION:
? OTHER INFORMATION:
? OTHER INFORMATION:
? OTHER INFORMATION:
? OTHER INFORMATION:
? OS-08-379-621-1

```

Query Match	11.2%;	Score 557.8;	DB 1;	Length 690;
Best Local Similarity	90.0%;	Pred. NO. 8.6e-165;		
Matches 620; Conservative	0;	Mismatches 42;	Indels 27;	Gaps 1;

QY	4090	TTGAGGGGGGTGTGGGCCACCCACCATTCCTGGCCCTGGCAATCGATGCACAAACA	41439
Db	2	TTGTAGGGGGGTGTGTGGCCACCCACCATTCCTGGCCCTGGCAATCGATGCACAAACA	61
QY	4150	GGTTTGGAAAGCACTTCTGTGGAGGCCCTTAATATCCCAAGTGGGTGTGATGCTGTG	42099
Db	62	GGTTTGGAAATGCACTTCTGTGGAGGCCACCTGTATATCCCAAGTGGGTGTGATGCTGTG	121
QY	4210	CTACAGCTTGAAGAAGTCTCTCAAGGCTTTCATCTCTACAAGTATCTCTGGGTGCACAC	42659
Db	122	CCACATGCTTGGGAAGTCTCCCAAGGCTTTCATCTCTACAAGTATCTCTGGGTGCACAC	181
QY	4270	AAGAAGTGAACCTCGAATCTCATATGTTCAAGAAATAGAATGTCTAGGCTGTCTTGGAGC	43299
Db	182	AAGAAGTGAATCTCGAATCGAATGTCACAGAAATAGAATGTCTAGGCTGTCTTGGAGC	241
QY	4330	CCACACACAGACATATTTGCTCTGTCTCTCAAGCTCAAGCAGGCTCTCGTATCATCTGCAAA	43899
Db	242	CCACACAGAAAGATATTTGCTCTGTCTCTCAAGCTCAAGCAGGCTCTCGTATCATCTGCAAA	301
QY	4390	TAAATGCAGCTTGTCTGCCATCCCGAGCTACATGATGTCACCGCCGAGACTGAATGTTCACA	44499
Db	302	TAAATGCCAGCTTGTCTGCCATCCCGAGCTACATGATGTCACCGCCGAGACTGAATGTTCACA	361
QY	4450	TCACATGGCTGGGAGAAACCCAAAGTACCTTTTGGAGTGGCTTCTCAAGAAAGCCAGC	45099
Db	362	TCACATGGCTGGGAGAAACCCAAAGTACCTTTTGGAGTGGCTTCTCAAGAAAGCCAGC	421
QY	4510	TCCCTGTATATGAGATGAAGTGTGCATCATATTAAGT-----	4548
Db	422	TCCCTGTATATGAGATGAAGTGTGCATCATATTAAGT-----	481
QY	4549	-----ATATTTGTGTGAGCATTTTGGCCAGAGGACGACAGTGTGCCAGGGTGCAGATG	4602
Db	482	CCACGCACTGTGTGTGGGCAATTTGGCCGAGGACACTGACAGTTGCCAGGGTGCAGATG	541
QY	4603	GAGGAGCTCTGCTTGTCTTCGAGAAAGCAATATCATTTTCAAGAGATCACTTCTTGGG	4662
Db	542	GAGGAGCTCTGCTTGTCTTCGAGAAAGCAATATCATTTTCAAGAGATCACTTCTTGGG	601
QY	4663	GTTTGGCTGTGAGAGCCCCCAATTAAGCCGTGGTCTATAGCCGTGTTCAAGGTTGTGA	4722
Db	602	GTTTGGCTGTGAGAGCCCCCAATTAAGCCGTGGTCTATAGCCGTGTTCAAGGTTGTGA	661
QY	4723	CTTGGATTGAGGAATGATGAGAAATAAT	4751
Db	662	CTTGGATTGAGGAATGATGAGAAATAAT	690

```

1  TITLE OF INVENTION:  INHIBITORS OF PLASMIN
2  NUMBER OF SEQUENCES:  10
3  CORRESPONDENCE ADDRESS:
4  ADDRESSEE:  HALE AND DORR LLP
5  STREET:  60 State Street
6  CITY:  Boston
7  STATE:  MA
8  COUNTRY:  United States of America
9  ZIP:  02109
10
11  COMPUTER READABLE FORM:
12  MEDIUM TYPE:  Floppy disk
13  COMPUTER:  IBM PC compatible
14  OPERATING SYSTEM:  PC-DOS/MS-DOS
15  SOFTWARE:  PatentIn Release #1.0, Version #1.30
16  CURRENT APPLICATION DATA:
17  APPLICATION NUMBER:  US/08/889,078
18  FILING DATE:  7 JULY 1997
19  CLASSIFICATION:  536
20  ATTORNEY/AGENT INFORMATION:
21  NAME:  Baker, Hollie L.
22  REGISTRATION NUMBER:  31,321
23  REFERENCE/DOCKET NUMBER:  102378.187CON
24  TELECOMMUNICATION INFORMATION:
25  TELEPHONE:  (617) 526-6000
26  TELEFAX:  (617) 526-5000
27  INFORMATION FOR SEQ. ID NO. 1:
28  SEQUENCE CHARACTERISTICS:
29  LENGTH:  690 base pairs
30  TYPE:  nucleic acid
31  STRANDEDNESS:  double
32  TOPOLOGY:  linear
33  MOLECULE TYPE:  cDNA
34  HYPOTHETICAL:  NO
35  ANTI-SENSE:  NO
36  ORIGINAL SOURCE:
37  ORGANISM:  Homo sapiens
38  FEATURE:
39  NAME/KEY:  CDS
40  LOCATION:  1..690
41  OTHER INFORMATION:
42  OTHER INFORMATION:  /partial
43  OTHER INFORMATION:  /codon_start=1
44  OTHER INFORMATION:  /function= "encodes plasmin protease domain"
45  OTHER INFORMATION:  /product= "nucleotide with corresponding
46  OTHER INFORMATION:  protein"
47  OTHER INFORMATION:  /number= 1
48
49  JS-08-889-078-1

```

	Query Match	11.2%	Score 557.8	DB 2	Length 690
	Best Local Similarity	90.0%	Pred. No. 8,66-165		
	Matches 620	Conservative 0	Mismatches 42	Indels 27	Gaps 1
QY	4090	TTGTAGGGGGGTGTGTGGCCACCCACATTCCTGGCCCTGGCAAGTCAGTCTGAGAACAA	4149		
Db	2	TTGTAGGGGGGTGTGTGGCCACCCACATTCCTGGCCCTGGCAAGTCAGTCTTGAACA	61		
QY	4150	GGTTTGAAGGACCTCTGTGAGGACACCTTAATATCCCAAGAGTGGTCTGACATGCG	4209		
Db	62	GGTTTGAATGCACCTCTGTGAGGACACCTTGATATCCCAAGTGGTGTACTGTCTG	121		
QY	4210	CTCAGTCTTGAAGAAGTCTCAAGGCTTCATCCTACAAAGTCAATCTGGGTGCACACC	4269		
Db	122	CCCACTGCTTGGAGAGTCCCAAGGCTTCATCCTACAAAGTCAATCTGGGTGCACACC	181		
QY	4270	AAGAAGTGAACCTCGGAATCTCATGTTCAGAGAAATGAAGGTCTAGAGCTTCTTGGAGC	4329		
Db	182	AAGAAGTGAATCTCGGAACCTCGATGTCCAGAAATGAAGGTCTTAGAGCTTCTTGGAGC	241		
QY	4330	CCACACAGACAGATATATGCTTGTCTTAAAGCTAACAGAGCTGCGTCACTACATGACAAG	4389		
Db	242	CCACAGAGAAAGATATATGCTTGTCTTAAAGCTAACAGAGCTGCGCTATCATCGACAAG	301		
QY	4390	TAAATCCAGCTGTCTGCCATCCCAAGACTACATGCTACCGCCAGGACTGAATGTTACA	4449		

```
Dh 302 TAATCCAGCTGTGTGCAATCCCAATTAATGTGTGGTGCAGCCGAGCCGAATGTTTCA 361
Qy 4450 TCACGTGCTGGGAGAAAACCAAGTACCTTTGGAGTGGCCCTTCTCAAGAACCCGAC 4509
    |||||
Db 362 TCACGTGCTGGGAGAAAACCAAGTACCTTTGGAGTGGCCCTTCTCAAGAACCCGAC 421
Qy 4510 TCCTGTATTAAGATGAAGTGTACATCACTATAAGT----- 4548
    |||||
Db 422 TCCTGTATTAAGATGAAGTGTACATCACTATAAGT----- 481
Qy 4549 -----ATATTTGTGCTGAGCATTTGGCCAGAGCCACTGACAGTTCAGAGGTGACAGTG 4602
    |||||
Db 482 CCACGGAATCTGTGTGGGCAATTTGGCCGAGGCATGACAGTTCAGAGGTGACAGTG 541
Qy 4603 GAGGGCTGTGTTGTTGCTTCGGAAGAGCAATATATTTTACAAGAGTCACTTTTGGG 4662
    |||||
Db 542 GAGGTCTCTGTGTTGCTTCGGAAGAGCAATATATTTTACAAGAGTCACTTTTGGG 601
Qy 4663 GTCTTGCTGTGCAGCCCAATTAAGCCGTGCTATGCTGTTTCAAGTTTGTTA 4722
    |||||
Db 602 GTCTTGCTGTGCAGCCCAATTAAGCCGTGCTATGCTGTTTCAAGTTTGTTA 661
Qy 4723 CTTGATTGAGGGAATGATGAGAAATAT 4751
    |||||
Db 662 CTTGATTGAGGGAATGATGAGAAATAT 690
```

```
RESULT 14
US-09-206-059-29
; Sequence 29, Application US/09206059
; Patent No. 6201104
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas
; TITLE OF INVENTION: Angiogenesis-Inhibiting Protein Binding Peptides and
; FILE REFERENCE: 05213-0370
; CURRENT APPLICATION NUMBER: US/09/206,059
; CURRENT FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-206-059-29
```

Query Match 5.1%; Score 257.2; DB 4; Length 1134;
Best Local Similarity 66.8%; Pred. No. 2.4e-70;
Matches 401; Conservative 0; Mismatches 163; Indels 36; Gaps 1;

```
Qy 3159 TATCGAGGCAATATCCACACACATGTCACAGGAAGCACTGTCATCTTGTCATCTATG 3218
    |||||
Db 571 TATCGAGGCAATATGCTGTACCGCTGTCCGCGACACCTGTGACGATGAGTGACAG 630
Qy 3219 ATACCACTGTCATGAGAGCAACCCCAAAATACCAAAATGTCGCTGACCGAGAAC 3278
    |||||
Db 631 ACCCTCAGACATTAACAGAGAACCAAAAATTTCCGCAAAAATTTGGATGAAAC 690
Qy 3279 TACTGACGATCCAAATTCCTGGGAAACAACCTGTGTTACCAACGATCCGTTGTG 3338
    |||||
Db 691 TACTGCGCAATCTGACGAAAAAGGCGCCATGCTGACATCAACCAAGCCAAAGTG 750
Qy 3339 AGTGGGAGTACGACATGACACAAATGCTCGAAGAAACGAATCAGTGTCCATAGAGCT 3398
    |||||
Db 751 CGGTGGAGTACTGTAAATACCGTCTG----- 779
Qy 3399 CCCACTGTGTTCCAGTTCACACATGAGGCTCATTTGAAACGACACCACTGAGCAA 3458
    |||||
Db 780 -----TGACTCCTCCCAATATCCAGGAACATTTGGCTCCACACACACTGAGCTA 834
Qy 3459 ACCCTGTGTCGCGAGTCTTACCATGTATAGCCAGATTAAGAGGACATTTCTCC 3518
    |||||
```

```
Dh 835 ACCCTGTGTCAGAGACTGCTACCATGTTGATGACAGAGTACGAGGACACATCTCC 894
Qy 3519 ACCACTGTCAGAGAGACATGTCATTTGGTTCATTCATGACACACACCGGCAATCAG 3578
    |||||
Db 895 ACCACACACAGAGAAAGAGTGTACGTGTTGTCATCTATGACACACACCGGCAAC 954
Qy 3579 AGACCCAGAAAACCTACCCAAATGATGCTGTCATGAATGAACTACTGACAGATCCAGAT 3638
    |||||
Db 955 AAGACCCAGAAAACCTACCCAAATGCTGCTGACATGAACTACTGACAGATCCAGAT 1014
Qy 3639 GCCGATACAGGCCCTGTGTTTACCATGAGACCCAGATCAGTGGAGTACTGCAAC 3698
    |||||
Db 1015 GCCGATAAAGCCCTGTGTTTACCATGAGACCCAGCTCAGTGGAGTACTGCAAC 1074
Qy 3699 CTGACGATGCTCAGACACAGAGGACATGTGCTGCTCTCCGATGTCATCCAGGTT 3758
    |||||
Db 1075 CTGAAAAATCTCAGGAACAGAACGAGTGTATGACACCTCGCCTGTGCTGCTT 1134
```

```
RESULT 15
US-08-030-410-2
; Sequence 2, Application US/08030410
; Patent No. 6221359
; GENERAL INFORMATION:
; APPLICANT: Komiyama, Atsushi
; APPLICANT: Nakahata, Tatsutoshi
; APPLICANT: Kubo, Tetsuo
; APPLICANT: Tanaka, Ryuhel
; APPLICANT: Kawano, Genji
; APPLICANT: Sano, Emiko
; APPLICANT: Kojima, Katsuki
; TITLE OF INVENTION: HEMOPOIETIC STEM CELL MULTIPLIER
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/030,410
; FILING DATE: 19930521
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digilio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8898
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2169
US-08-030-410-2
```

Query Match 1.8%; Score 90.8; DB 4; Length 2172;
Best Local Similarity 50.5%; Pred. No. 5.7e-18;
Matches 279; Conservative 0; Mismatches 237; Indels 36; Gaps 1;

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 06:40:43 ; Search time 758 Seconds
(without alignments)
4405.804 Million cell updates/sec

Title: US-09-923-515-3_COPY_8939_13938

Perfect score: 5000

Sequence: 1 ctcgcactgttaccgccggtt.....ctgcattatttgatttga 5000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 478924 segs, 333959956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/PCITUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3753.2	75.1	10422	9	US-09-870-759-125
2	1106.4	22.1	2497	10	US-09-946-893-1
3	654.6	13.1	1907	9	US-09-924-340-53
4	654.6	13.1	1907	9	US-09-992-600A-53
5	306.4	6.1	1236	10	US-09-880-107-2341
6	257.2	5.1	1134	10	US-09-873-676-29
7	235.6	4.7	384	10	US-09-960-352-10993
8	231.4	4.6	468	10	US-09-960-352-10851
9	229.2	4.6	372	10	US-09-960-352-7521
10	204.4	4.1	313	9	US-09-796-692-5593
11	197	3.9	394	10	US-09-960-352-11088
12	193.8	3.9	346	10	US-09-960-352-12743
13	145.2	2.9	422	10	US-09-833-381-252
14	139.6	2.8	234	10	US-09-960-352-14383
15	134.8	2.7	199	10	US-09-960-352-11543
16	134.8	2.7	431	10	US-09-864-761-13941
17	129.6	2.6	160	10	US-09-864-761-20992
18	122.4	2.4	408	10	US-09-960-352-1535
19	108.4	2.2	203	10	US-09-960-352-11298

20	104.2	2.1	168	10	US-09-864-761-30505	Sequence 30505, A
21	94	1.9	362	10	US-09-864-761-4229	Sequence 4239, Ap
22	93.4	1.9	242	10	US-09-960-352-12971	Sequence 12971, A
23	89.2	1.8	1344	9	US-10-081-309-1	Sequence 1, Appl1
24	89.2	1.8	5898	10	US-09-880-107-3708	Sequence 3708, Ap
25	80.4	1.6	645	12	US-10-036-869-22	Sequence 22, Appl
26	80.4	1.6	1284	12	US-10-036-869-24	Sequence 24, Appl
27	77.2	1.5	780	9	US-09-870-759-96	Sequence 96, Appl
28	71.6	1.4	402	10	US-09-960-352-15108	Sequence 15108, A
29	68	1.4	792	10	US-09-888-615-37	Sequence 37, Appl
30	67.6	1.4	614	10	US-09-879-792-33	Sequence 33, Appl
31	66.6	1.3	1222	9	US-10-125-459-3	Sequence 3, Appl1
32	66.6	1.3	1222	9	US-10-067-761-5	Sequence 5, Appl1
33	66.6	1.3	1222	10	US-09-946-633-5	Sequence 5, Appl1
34	66.6	1.3	1222	10	US-09-879-792-35	Sequence 35, Appl
35	66	1.3	1230	10	US-09-888-615-45	Sequence 45, Appl
36	66	1.3	1614	10	US-09-879-792-11	Sequence 11, Appl
37	66	1.3	1079	9	US-10-125-459-5	Sequence 5, Appl1
38	64.6	1.3	1079	9	US-10-067-761-7	Sequence 7, Appl1
39	64.6	1.3	1079	10	US-09-804-156-7	Sequence 7, Appl1
40	64.6	1.3	1079	10	US-09-946-633-5	Sequence 5, Appl1
41	64.6	1.3	1079	10	US-10-125-459-4	Sequence 4, Appl1
42	64.6	1.3	1394	9	US-10-067-761-6	Sequence 6, Appl1
43	64.6	1.3	1394	10	US-09-804-156-6	Sequence 6, Appl1
44	64.6	1.3	1394	10	US-09-946-633-4	Sequence 4, Appl1
45	64.6	1.3	1394	10	US-09-946-633-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-870-759-125
; Sequence 125, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIORITY FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 125
; LENGTH: 10422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6667)..(10176)
; OTHER INFORMATION:
US-09-870-759-125

Query Match 75.1%; Score 3753.2; DB 9; Length 10422;
Best local Similarity 99.8%; Pred. No. 0;
Matches 3758; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1235 TGGCTTGATGATGACTACTGAGGATCCAGATCCTGTGCGAGCCCTTATGTTATAC 1294
|||||
Db 6657 TGGCTTGATGATGACTACTGAGGATCCAGATCCTGTGCGAGCCCTTATGTTATAC 6716
|||||
QY 1295 GAGGATGCCAGTGGAGTGGAGATGCAACCTGCACATGCTCAGACGAGAGG 1354
|||||
Db 6717 GAGGATGCCAGTGGAGTGGAGATGCAACCTGCACATGCTCAGACGAGAGG 6776
|||||
QY 1355 GACTGCCGTGCGCCCTCCACTATTACCCGATTCACAGCCCTAGAGGCTCTTGGAACA 1414
|||||
Db 6777 GACTGCCGTGCGCCCTCCAGCTGTACCCCGGTTCACAGCCCTAGAGGCTCTTGGAACA 6836
|||||
QY 1415 AGACCACTGAGCAAGGCTGGGGTGGAGAGTGCTACCGAGGAATGACAGATTA 1474
|||||

QY 3635 AGATCCGATACAGAGCCCTTGCTGTTTTACATGAGACCCAGCATAGGTGGAGTACTG 3694
|||||
Db 9057 AGATGCCATACAGGCCCTTGCTGTTTTACATGAGACCCAGCATAGGTGGAGTACTG 9116
QY 3695 CAACCTGACGGATGCTCAGACACAGAAAGGACTGTGCTCGCTCTCCGACTGTCATCA 3754
|||||
Db 9117 CAACCTGACGGATGCTCAGACACAGAAAGGACTGTGCTCGCTCTCCGACTGTCATCA 9176
QY 3755 GGTTCGAAGCCCTAGGGCCCTTCTGACAAACGATGATGTTGGGAATGGGAAGATA 3814
|||||
Db 9177 GGTTCGAAGCCCTAGGGCCCTTCTGACAAACGATGATGTTGGGAATGGGAAGATA 9236
QY 3815 CCGGGGCAAGAGGCAACACACTGTACTGTGGAGCGCCATGCCAGATGGGCGCCAGGA 3874
|||||
Db 9237 CCGGGGCAAGAGGCAACACACTGTACTGTGGAGCGCCATGCCAGATGGGCGCCAGGA 9296
QY 3875 GCGCCATACACAGACAGACGTCATTCCAGGACAATAAATGGGAGGTCTGGAAAAAA 3934
|||||
Db 9297 GCGCCATACACAGACAGACGTCATTCCAGGACAATAAATGGGAGGTCTGGAAAAAA 9356
QY 3935 TTACTGCCCTAACCTGATGATGATCATCATATGTCCTCGTGTCTACACAAATATCCAG 3994
|||||
Db 9357 TTACTGCCCTAACCTGATGATGATCATCATATGTCCTCGTGTCTACACAAATATCCAG 9416
QY 3995 AAAAATTTTGACTACTGTGATATCCCTCTCTGTGCATCCCTTCATTGTTGATGGGA 4054
|||||
Db 9417 AAAAATTTTGACTACTGTGATATCCCTCTCTGTGCATCCCTTCATTGTTGATGGGA 9476
QY 4055 GGCCTCAAGTGAAGCCGAGAAAATGCTCGGAAGCATTTGAGGGGGTGTGGCCACCC 4114
|||||
Db 9477 GGCCTCAAGTGAAGCCGAGAAAATGCTCGGAAGCATTTGAGGGGGTGTGGCCACCC 9536
QY 4115 ACATTCCTGGCCCTGGCAAGTCAGTCCTGCAACAAGGTTTGGAAAGCACTTCTGTGGAG 4174
|||||
Db 9537 ACATTCCTGGCCCTGGCAAGTCAGTCCTGCAACAAGGTTTGGAAAGCACTTCTGTGGAG 9596
QY 4175 CACCTTAATATCCCGAGATGGGTGCTGACTGCTCACTGCTTGAGAAAGTCCCTCAAG 4234
|||||
Db 9597 CACCTTAATATCCCGAGATGGGTGCTGACTGCTCACTGCTTGAGAAAGTCCCTCAAG 9656
QY 4235 GGCCTTCATCCCTCAAGGATATCCCTGGGTGACACACAGAAAGTGAACCTCGAATCTCATGT 4294
|||||
Db 9657 GGCCTTCATCCCTCAAGGATATCCCTGGGTGACACACAGAAAGTGAACCTCGAATCTCATGT 9716
QY 4295 TCGAAGAAATAGAGTGTCTAGGGCTGTTCTGGAGCCACACAGAGATATGGCTGGCT 4354
|||||
Db 9717 TCGAAGAAATAGAGTGTCTAGGGCTGTTCTGGAGCCACACAGAGATATGGCTGGCT 9776
QY 4355 AAAGCTAAGAGGCTGCGCTGATCATCTAGTCAAAAGTATGCGAGCTTGTCTGCCATCCCC 4414
|||||
Db 9777 AAAGCTAAGAGGCTGCGCTGATCATCTAGTCAAAAGTATGCGAGCTTGTCTGCCATCCCC 9836
QY 4415 AGACTACATGATGTCACCGCCAGACTCAATGTTACATCACTGGCTGGGGGAAACCCAGG 4474
|||||
Db 9837 AGACTACATGATGTCACCGCCAGACTCAATGTTACATCACTGGCTGGGGGAAACCCAGG 9896
QY 4475 TACCTTGGGACTGGCTTCTCAAGGAAGCCAGCTCTTGTATTGAGAATGAAGTGTG 4534
|||||
Db 9897 TACCTTGGGACTGGCTTCTCAAGGAAGCCAGCTCTTGTATTGAGAATGAAGTGTG 9956
QY 4535 CAATCACTATAGATATTTGCTGCTAGCATTTGGCCAGAGGCACTGACAGTTGCCAGGG 4594
|||||
Db 9957 CAATCACTATAGATATTTGCTGCTAGCATTTGGCCAGAGGCACTGACAGTTGCCAGGG 10016
QY 4595 TGACACTGGAGGGCTCTGCTTGGCTTGCAGAAAGGACAAATACATTTTCAAGAGACTCAC 4654
|||||
Db 10017 TGACACTGGAGGGCTCTGCTTGGCTTGCAGAAAGGACAAATACATTTTCAAGAGACTCAC 10076
QY 4655 TTCTTGGGGCTTGGCTGTGACAGCCCAATAAGCTGTGTCATGCTGTTTCAAG 4714
|||||
Db 10077 TTCTTGGGGCTTGGCTGTGACAGCCCAATAAGCTGTGTCATGCTGTTTCAAG 10136
QY 4715 GTTTGTTACTTGATTTGAGGGAATGATGAGAAATAATTAATTGGACGGGAGACAGAGTGA 4774

Db 10137 GTTTGTTACTTGGATGAGGGAATGATGAGAAATAATTAATTGAGCGGGAGACAGAGTGA 10196
QY 4775 AGCATCAACCTACTTAAAGCTGGAACCTGGGTAAGGATTTAGCATGCTGGAATAATAG 4834
|||||
Db 10197 AGCATCAACCTACTTAAAGCTGGAACCTGGGTAAGGATTTAGCATGCTGGAATAATAG 10256
QY 4835 ACACAAATCAACGAGACACTGTTCACACTACACAGCTATGCGAAACCTTGGCATTTT 4894
|||||
Db 10257 ACACAAATCAACGAGACACTGTTCACACTACACAGCTATGCGAAACCTTGGCATTTT 10316
QY 4895 GGTATTTTGTATTAAGCTTTTAAGTCTGACTGACAAATCTGTTAAGGTGCATA 4954
|||||
Db 10317 GGTATTTTGTATTAAGCTTTTAAGTCTGACTGACAAATCTGTTAAGGTGCATA 10376
QY 4955 GCTATGACATTTGTTAAAAAATAACTGCACTTATTTGATTTGA 5000
|||||
Db 10377 GCTATGACATTTGTTAAAAAATAACTGCACTTATTTGATTTGA 10422

RESULT 2
US-09-946-893-1
; Sequence 1, Application US/09946893
; Patent No. US20020072494A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
; TITLE OF INVENTION: Inhibitors
; FILE REFERENCE: Newburn
; CURRENT APPLICATION NUMBER: US/09/946, 893
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,893
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentl Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(2482)
; US-09-946-893-1

Query Match 22.1%; Score 1106.4; DB 10; Length 2497;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1334; Conservative 0; Mismatches 241; Indels 63; Gaps 2;

QY 3159 TATCGAGGATATCTCCACCACTGTCTACAGGAAGACCTGTCAATCTTGTCATCTATG 3218
|||||
Db 896 TATCGGGGAATGTGCTGTACCGTGTCCGGCACACCTGTCTAGCACTGGAGTGCACAG 955
QY 3219 ATACCACTGGCATCAGAGACCCCGAAGAACTACCAAAATGCTGGCCGTGACGAGAAC 3278
|||||
Db 956 ACCCTCTACACATTAACAGACACCAAGAACTTCCCTGCAAAAATTTGATGAATAAC 1015
QY 3279 TACTCAGAAATCCAGATTCTGGGAAACACCTGTGTGTACACACGATCCCTGTGTG 3338
|||||
Db 1016 TACTCAGAAATCCAGATTCTGGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1075
QY 3339 AGGTGGAGTACTGCAATCTGACACATGCTCAGAAACGATCAGATGCTTGTGAGACT 3398
|||||
Db 1076 CGGTGGAGTACTGCAATCTGACACATGCTCAGAAACGATCAGATGCTTGTGAGACT 1104
QY 3399 CCCAGTGTGTTCCAGTTCGAAGGCTCATTTCTGAAGAGCACCACACTGAGCAA 3458
|||||
Db 1105 -----TGACTCTCTCCCGATATCCAGGAAATTTGGTCTCCACAGCACACTGAGCTA 1159
QY 3459 ACCCTGTGTGTCGGCAGTGTACCATGTAATGGCCAGAGTTTCAGAGGACATTTCTCC 3518
|||||
Db 1160 ACCCTGTGTGTCAGAGTGTGTACCATGTAATGGCCAGAGTTTCAGAGGACATTTCTCC 1219
QY 3519 ACCACTGTACAGGAAGACATGTCAATCTTGTGTCATGACACACACCGGACATCAG 3578

```
Db 1220 ACCACACACAGAGAAAGATGTCACTCTTGATCATCTATGACACACACCGGCACAG 1279
QY AGAACCCAGAAAATCTACCCAAATGATGCCCTGACAAATGAATCTACTCAGAAATCCAGAT 3638
Db 1280 AAGACCCAGAAAATCTACCCAAATGATGCCCTGACAAATGAATCTACTCAGAAATCCAGAT 1339
QY GCGATACAGGCGCTTGCTGTTTACATGAGACCCCAAGCATCAGAGTGGAGTACTGCAAC 3698
Db 1340 GCGCATAAAGCGCCCTGCTGTTTACACAGACCCCAAGCATCAGAGTGGAGTACTGCAAC 1399
QY CTGACGGGATGCTCAGACACAGAAAGGACTGCTGCTCCGCTCCGATGATCATCAGATT 3758
Db 1400 CTGAAAAATGCTCAGAGAACGAGAGTGTGTACACACTCCGCTGTGTCTGCTT 1459
QY 3759 CCAAGCTAGGGCTCTCTTGAAACAAGATGATGTTTGGGAATGGGAAAGATACCGG 3818
Db 1460 CCAAGATGTAGAGACTCTTCCGAAAGAACTGATGTTTGGGAATGGGAAAGATACCGA 1519
QY 3819 GGCAGAGAGCAACACTGTTACTGAGAGCCATGCCAGGAATGGGCTGCCAGAGGCC 3878
Db 1520 GGCAGAGAGGCGCACTGTTACTGAGAGCCATGCCAGAGACTGGCTGCCAGAGGCC 1579
QY 3879 CATGACACAGCAGCTTCATTCAGGACAAATGAATGGGAGGCTCGAAAAAATTAC 3938
Db 1580 CATGACACAGCAGCTTCATTCAGGACAAATCCAGGCGGGCTCGAAAAAATTAC 1639
QY 3939 TGCGGTAAACCTGATGGTGAATCAATGATGCCGCTGCTACCAATGATCCAGAAAA 3998
Db 1640 TGCGGTAAACCTGATGGTGAATGATGCTGCTGCTGCTACCAATGATCCAGAAAA 1699
QY 3999 CTTTGTACTACTGATGATCCCTCTCTGTGATCCTCTTCAATTTGATGGGAAAGCT 4058
Db 1700 CTTTGTACTACTGATGATCCCTCTCTGTGATGCTGCTGCTTCAATTTGATGGGAAAGCT 1759
QY 4059 CAATGAGAGCGGAGAAATGTCCTGGAGAGATGTAGGGGGGTGTGTGGCCACCCACAT 4118
Db 1760 CAATGAGAGCGGAGAAATGTCCTGGAGAGGTTGTAGGGGGGTGTGTGGCCACCCACAT 1819
QY 4119 TCCTGGGCGCTGGCAAGTCACTCTCAGAAACAAGTTTGGAAAGCACTCTGTGGAGGACC 4178
Db 1820 TCCTGGGCGCTGGCAAGTCACTCTTGAACAAGTTTGGAAATGCACTCTGTGGAGGACC 1879
QY 4179 TTAATATCCCAAGATGGGTGTGACTGTCTCAGTCTCTGGAAGAGTCCCTCAAGGCT 4238
Db 1880 TTAATATCCCAAGATGGGTGTGACTGTCTCAGTCTCTGGAAGAGTCCCTCAAGGCT 1939
QY 4239 TCATCTACAAAGTCACTCCGCGGTGGACACCAAGAAAGTGAATCCGATCTCATGTTCA 4298
Db 1940 TCATCTACAAAGTCACTCCGCGGTGGACACCAAGAAAGTGAATCTCGAACCGCATGTTCA 1999
QY 4299 GAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACAAACAGATATTTGCTTGTCTAAAG 4358
Db 2000 GAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACAAAGAAATATTTGCTTGTCTAAAG 2059
QY 4359 CTAAAGAGGCGCTGCTCATCTACACAAAGTAATGCCAGCTTGTCTCCATCCCAAGAC 4418
Db 2060 CTAAAGAGGCGCTGCTCATCTACACAAAGTAATCCCAAGCTTGTCTCCATCCCAAGAC 2119
QY 4419 TACATGCTACCGCGAGAGTGAATGTATCATCTAGCTGGGGGAGAAACCAAGGTACC 4478
Db 2120 TATGTGCTGCTGACCGGACCGAATGTTTCATACGTGGGGGAGAAACCAAGGTACT 2179
QY 4479 TTTGGAGCTGGCTTCTCAAGAGAGCCAGCTCTTGTATTGAGATGAGATGTCAT 4538
Db 2180 TTTGGAGCTGGCTTCTCAAGAGAGCCAGCTCTCTGTATTGAGATGAGATGTCAT 2239
QY 4539 CACTATAAGT-----AATTTGCTGCTGAGCAATTTGGCC 4571
Db 2240 CGCATATGATTTCTGAATGGAAGATCCATCCACCGAACTGTGTCTGGGCAATTTGGCC 2299
QY 4572 AGAGGACTACAGTTGCCAGGGTGACAGTGGAGGGCTCTGTGTTGCTTCCAGAAAGAC 4631
```

```
Db 2300 GGAAGCATGACAGTTGCCAGGGTGACAGTGGAGTCTCTGTGTTGCTTGAGAGAGAC 2359
QY 4632 AATATATTTTACAAAGAGTCACTTCTTGGGCTTGTGGCTGTGACAGCCCAATTAAGCT 4691
Db 2360 AATATATTTTACAAAGAGTCACTTCTTGGGCTTGTGGCTGTGACAGCCCAATTAAGCT 2419
QY 4692 GGTGCTATGCTGCTGTTTCAAGGTTTGTACTTGTGATTTGAGGGAATATAGAAATAT 4751
Db 2420 GGTGCTATGCTGCTGTTTCAAGGTTTGTACTTGTGATTTGAGGGAATATAGAAATAT 2479
QY 4752 TAATTGACGGGAGACAG 4769
Db 2480 TAATTGACGGGAGACAG 2497
```

RESULT 3
US-09-924-340-53
Sequence 53, Application US/09924340
Publication No. US20030027248A1

GENERAL INFORMATION:
APPLICANT: Bejatin, Stephane
APPLICANT: Tanaka, Hiroaki
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
FILE REFERENCE: 91. US2. REG
CURRENT APPLICATION NUMBER: US/09/924, 340
CURRENT FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: US 60/305,456
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/302,277
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/298,698
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/293,574
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 112
SOFTWARE: JPlatent
SEQ ID NO 53
LENGTH: 1907
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..1043
NAME/KEY: CDS
LOCATION: 1044..1664
NAME/KEY: 3'UTR
LOCATION: 1665..1907
NAME/KEY: POLYA_Signal
LOCATION: 1869..1874
NAME/KEY: POLYA_site
LOCATION: 1892..1907

US-09-924-340-53

Query Match 13.1%; Score 654.6; DB 9; Length 1907;
Best Local Similarity 86.2%; Pred. No. 5e-196;
Matches 766; Conservative 0; Mismatches 79; Indels 44; Gaps 2;

```
QY 4132 AAGTAGCTGTGAGAACGAGTTTGGAAAGCACTTGTGGAGGACCTTAATATCCAG 4191
Db 1018 AACTTAATCTCTTTCCAGGTTTGGAAATGCACTTGTGGAGGACCTTGAATATCCAG 1077
QY 4192 AGTGGTGTGCTAGCTCTCTCATCTGTAAGAAGTCTCTCAAGGCTTTCATCTACAG 4251
Db 1078 AGTGGTGTGCTAGCTCTCTCCACCTGCTTGGAGAGTCTCCCAAGGCTTTCATCTACAG 1137
QY 4252 TCATCTGTGGTGCACACCAAGATGAATCTCATGTTTCATGAGAAATGAGAGT 4311
Db 1138 TCATCTGTGGTGCACACCAAGATGAATCTGAAACGCAATGTTCAAGAAATGAGAGT 1197
QY 4312 CTAGGCTGTCTTGGAGGCCACACAGAGATATTTGCTGCTTAAGCTAAGGCGCTG 4371
Db 1198 CTAGGCTGTCTTGGAGGCCACACAGAAATATTTGCTGCTTAAGCTAAGGCGCTG 1257
```


Db	121	TCGTGAGGATGATTCAAACGGGCCC	CGTGTCTACACCATGAAACCCAAAGAACCTTTT	TGA	180
QY	4007	CTATGTGATATCCCTCTCTGTGCATCCTCTTATTTGATTGTGGAGCCCTGAAGTGA		4066	
Db	181	CTACTGTAGCTCCCTCCTAGTGTAGTCC	--TCATTGCACTGTGGAAACCCAAAGTGA		237
QY	4067	GCCGAAATAATGTCTGTGAAGCATTTGAGGGGGGTGTGGGCCACCCACATTTCCGGCC		4126	
Db	238	GCCGAAATAATGTTCTGTGAAGGATTTGACTGGGTGTGTCTCCAAACCCACTCTTGGCC		297	
QY	4127	CTGGCAAGTCAGTCTCGAACCAAGTTTGGAAAGCACTTCTGTGAGGACCCTTAATTC		4186	
Db	298	CTGGCAAGTCAGGCTTCGACACAAAGCT	---GAGGACACTCTGTGGGAGAACTCATATTC		354
QY	4187	CCCGAGTGGGTGCTGAC	4204		
Db	355	CCCGAAGTGGGTGCTGAC	372		

```

RESULT 10
US-09-796-692-5593
; Sequence 5593, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Maillon, Jane A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077_001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5593
; LENGTH: 313
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-5593

Query Match          4.1%; Score 204.4; DB 9; Length 313;
Best Local Similarity 89.4%; Pred. No. 3,7e-54;
Matches 220; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Oy   3906 ACAATAATATGGCAGGCTCGAAAAAAATTACTGCCGTACCCGTGATGCATCAAT 3965
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    68 AAAAACCAAACCGCATGTGTTCACACAGTACGTCGCGTAACCTGATGGTGATCAAT 127

3966 GGTCCTGGTGCTCACACATGAATCCAGAAGAACTTTTGACTACTGTGATATCCCTCTC 4025

```

Accession	Sequence	Position
Db	128 GGTCCCTGTCTACAGAGTGAATCCAGAAACTTTTGACTAGTATATCCCTTC	187
Oy	4026 TGTGCATCTCTTCAATTGATTTGGGAACCCCAAGTGGACCGAAGAAAGCTCTGA	4085
Db	188 TGTGCATCTCTTCAATTGATTTGGGAACCCCAAGTGGACCGAAGAAAGTCTCAGA	247
Oy	4086 AGCATTTGAGGGGGTGTGGCCCAACCAATTCCTGGCCCTTGGCAAGCTCAGTCA	4145
Db	248 AGCATTTGAGGGGGTGTGTGGCCCAACCAATTCCTGGCCCTTGGCAAGCTTACA	307
Oy	4146 ACAAGG 4151	
Db	308 ACAAGG 313	

```

RESULT 11
US-09-960-352-11088
: Sequence 11088, Application US/09960352
: Patent No. US20020137139A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Ningbing
: APPLICANT: Byatt, John C.
: APPLICANT: Mathialagan, Nagappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
: FILE REFERENCE: 16511.006/37-21(10298)C
: CURRENT APPLICATION NUMBER: US/09/960.352
: CURRENT FILING DATE: 2001-09-24
: NUMBER OF SEQ ID NOS: 15112
: SEQ ID NO 11088
: LENGTH: 394
: TYPE: DNA
: ORGANISM: Bos taurus
: OTHER INFORMATION: Clone ID: 47-LIB34-084-Q1-E1-D4
US-09-960-352-11088

```

Query Match	3.9%	Score 197	DB 10	Length 394
Best Local Similarity	72.0%	Pred. 9.6e-52		
Matches 283	Conservative 0	Mismatches 85	Indels 25	Gaps 1
QY 4248	AAGGTCATCCTGGGTGCACACCACAGAAAGTGAACCTCGAATCTCATGATCTTCACGAAATAGAA			
Db 2	AAGGACATCTGGTGTGCACACAAACGAAAGATCGCGGACACAGAGTGTCCAGGAAATACAA			
QY 4308	GTCGTAGGCTGTCTTCTTGGAGCCACACAGAGATATTCCTGCTTAAAGCTAAGCAGG			
Db 62	GTGTCCAGGCTGTTCGCGGAGCCCTCTTAGCGGACATTCGCTTACTCAAGCTGACACGA			
QY 4368	CTGTGCGTCATCATCTGACAAAGTAATGCGAGCTGTCTGCATCCGCCAGCATCATGTGC			
Db 122	CTGTGCATCATCATCGAAAGAGGTATCCAGCTGTGTGCGACCCCGCACTCATATGTGT			
QY 4428	ACCGCAGGACTGAATGTTATCATCTGCTGGGGAGAAACCCAAAGGTACTTGTGGAGCT			
Db 182	GCAGCCCGGACAGAAATGTTACTATCTGTGGGAGAAACCCAAAGCATTTTGGTGA			
QY 4488	GGCCTTCTCAAGGAAGCCAGCTCCTGTATTGAGAAATGAAGTGTCAATCACTATAAG			
Db 242	GGCCTCTCTTAAGGAAGACGACCTGCTGTGATCGAAGACAGTGTTATTCCTCAACAG			
QY 4548	TATATTGTGCTGAG-----CATTTGGCCAGAGGCACGA			
Db 302	TATCTGCAGCGGAAGGTCAAGCCACGAGCTGTGTGCGGCCATCTGTATTGGAAGGACTGA			
QY 4583	CAGTTGCCAGGGGTGACAGTGAAGGGCCCTCTGT			
Db 362	CAGCTGCCAAGGTGACAGAGGGGCGCTCTGT			

Thu Mar 6 17:02:00 2003

us-09-923-515-3_8939_13938.rnpb

Page 10

```

? CURRENT FILING DATE: 2001-09-24
? NUMBER OF SEQ ID NOS: 15112
? SEQ ID NO 11543
? LENGTH: 199
? TYPE: DNA
? ORGANISM: Bos taurus
? OTHER INFORMATION: Clone ID: 49-LIB34-068-Q1-E1-E10
? US-09-0660-352-11543

```

Query Match	2.7%;	Score 134.8;	DB 10;	Length 199;
Best Local Similarity	82.7%;	Pred. NO. 2.8e-32;		
Matches 167;	Conservative	0;	Mismatches 32;	Indels 3;
				Gaps 1;

QY	3926	GGAATAAAATCTAGCCGTAACCCATTGGTACATCAATAGTGCCCTGTGTACACAAT	3985
Db	1	GGAGAGAAACTACTGCCGTATCCTTAGCGTATGTCAAGGGCCCTGTGTACCAT	60
QY	3986	GAATCCAGAAGAACTTTTGTGACTACTGTATATACCCTCTGTGCATCCTTCATTTGA	4045
Db	61	GAOCCAAGSAGACTTTTTGTACTAGTACTGACGTCCTCCTAGTGTAGTGC---TCAITGGA	117
QY	4046	TTTGTGGAGACCTCAAGTGGAGGCCGAGAATAATGCTCTGGAGACATTTGAGGGGGGTGT	4105
Db	118	CTGTGGAAAACCCAAAGTGGAGCCGGAAGAAATGTTCTGGAAGATTTGAGTGGGTGT	177
QY	4106	GGGCCACCCACATCTCCGGGCC	4127
Db	178	GTCCAAAGCCACACTTTGGCCC	199

Search completed: March 6, 2003, 15:53:15
Job time : 901 secs

Gencore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 05:15:44 ; Search time 5999.5 Seconds
(without alignments)
13497.364 Million cell updates/sec

Title: US-09-923-515-3_COPY_8939_13938

Perfect score: 5000
Sequence: 1 ctcgcacatttcacccggtt.....ctgcacattatttgattga 5000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pla:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	606	12.1	856	9	AL579900 AL579900
2	601.4	12.0	766	13	BI759134 603042546
3	568	11.4	568	12	BF970104 602272382
C 4	546.8	10.9	766	10	AV700701 AV700701
5	502	10.0	859	12	BG402207 602465764
6	461.4	9.2	788	10	AV695776 AV695776

C 7	441.4	8.8	622	9	AI061613
8	396	7.9	762	12	BG428018
9	395.8	7.9	750	10	AV695796
C 10	392	7.8	593	10	BE348267
11	387.4	7.7	701	13	BI219735
12	383.4	7.7	611	10	AV662084
13	370	7.4	481	14	H60805
C 14	368.6	7.4	563	10	AM271976
15	365	7.3	758	12	BG616686
16	364.2	7.3	668	12	BG427320
17	361.6	7.2	594	10	AV661991
C 18	359.2	7.2	558	10	AM104579
C 19	357	7.1	553	9	AI633979
C 20	355.4	7.1	466	9	AI948806
21	350	7.0	504	14	N77239
C 22	348.2	7.0	541	9	AA722885
23	348.2	7.0	740	10	AV653775
C 24	346.8	6.9	636	10	AV645329
C 25	346.4	6.9	515	9	AI351543
C 26	345.4	6.9	568	14	T62069
C 27	341.2	6.8	462	9	AI768449
C 28	340.8	6.8	545	10	AV720081
C 29	340.8	6.8	621	9	AI377474
C 30	340.6	6.8	464	10	BE326689
31	340.6	6.8	595	10	AM950594
C 32	339	6.8	457	9	AI631756
C 33	334.6	6.7	462	9	AA995929
C 34	333.8	6.7	444	9	AI640315
C 35	332.6	6.7	483	10	AV720676
36	328.6	6.6	607	10	AV662061
C 37	325.8	6.5	466	9	AI531542
38	325.6	6.5	755	13	BI145879
39	324.8	6.5	614	10	AV693554
40	321.2	6.4	586	9	AA106793
41	319.6	6.4	424	14	H73861
42	319.2	6.4	687	10	AV650465
43	318.2	6.4	694	9	AA032930
44	318	6.4	564	10	BE65711
C 45	318	6.4	668	14	BQ598303

ALIGNMENTS

RESULT 1	AL579900/c	856 bp	MRNA	linear	EST 16-FEB-2001
LOCUS	AL579900				
DEFINITION	AL579900 LTI_NFL008.TC2 Homo sapiens cDNA clone CSDD002YJ21 3				
ACCESSION	AL579900				
VERSION	AL579900.1	GI:12945394			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.				
TITLE	1 (bases 1 to 856)				
JOURNAL	Full-length cDNA libraries and normalization				
COMMENT	Unpublished (2001)				
CONTACT	Genoscope				
Genoscope - Centre National de Sequencage					
BP 191 91006 EVRY cedex - France					
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.					
FEATURES	location/Qualifiers				
source	1. 856				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="CSDD002YJ21"				
	/clone_lib="LTI_NFL008.TC2"				
	/sex="male"				
	/tissue.type="T cells from T cell leukemia"				
	/note="Vector: pcMVSPORT 6; Site_1: NotI; 1st strand cDNA"				

Chinese National Human Genome Center at Shanghai
331 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

SOURCE

1. 766
Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="GKCAEH09"

/clone_lib="GKC"

/tissue_type="hepatocellular carcinoma"

/dev_stage="Adult"

/lab_host="SOLR"

/note="Vector: pBluescript sk(+); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 199 a 204 c 166 g 193 t 4 others

ORIGIN

Query Match

Best Local Similarity 84.0%; Score 546.8; DB 10; Length 766;
Pred. No. 3,8e-152;
Matches 657; Conservative 0; Mismatches 81; Indels 44; Gaps 2;

OY 4214 CTGCTTGAAGAGCTCCAGGCGCTGATCTACAGAGTCACTGGGTCACACCAAGA 4273
DB 765 CTGCTTGAAGAGCTCCAGGCGCTTATCTCTCAAGGTCATCTGGGTCACACCAAGA 706
OY 4274 AGTGAACCTCGAATTCATCTGTCAGAAATAGAAGTGTCTAGGCTGTTCTGGAGCCAC 4333
DB 705 AGTGAACCTCGAATTCATCTGTCAGAAATAGAAGTGTCTAGGCTGTTCTGGAGCCAC 646
OY 4334 ACAAGCAGATTTTCCTCTGCTTAAAGCTAAGCAGCCCTGCTCATCAGTACAAAGTAAT 4393
DB 645 ACGAAAGATATTTGCTGCTTAAAGCTAAGCAGCCCTGCTCATCAGTACAAAGTAAT 586
OY 4394 GCCAGCTTGTCCATCCCGACAGTACATGTCACCCGACGAGCTGATGTACATCAC 4453
DB 585 CGCAGCTTGTCCATCCCGACAGTACATGTCACCCGACGAGCTGATGTACATCAC 526
OY 4454 TGGCTGGGAGAAACCCAGGTCACCTTTGGAGTGGCTTCTCAAGAGCCAGCTCCT 4513
DB 535 TGGCTGGGAGAAACCCAGGTCACCTTTGGAGTGGCTTCTCAAGAGCCAGCTCCTC 466
OY 4514 TGTATTGAGATGAAGTGTGCATCACTCACTAAGT----- 4548
DB 465 TGTATTGAGATGAAGTGTGCATCACTCACTAAGT----- 4548
OY 4549 --ATATTGTGCTGACATTTGGCCAGAGCAGTGTCCAGGTCAGTGTGAGG 4606
DB 405 CGAAGCTGTCTGGGATTTGGCCAGAGCAGTGTCCAGGTCAGTGTGAGG 346
OY 4607 GCGCTGGTGTCTTCGAGAGGACAATATCTTTCAAGAGTCACTCTGGGGTCT 4666
DB 345 GCGCTGGTGTCTTCGAGAGGACAATATCTTTCAAGAGTCACTCTGGGGTCT 286
OY 4667 TGGCTGTGACGCCCCCAATTAAGCTGTCTATGCTGCTTTCAAGGTTTGTACTTG 4726
DB 285 TGGCTGTGACGCCCCCAATTAAGCTGTCTATGCTGCTTTCAAGGTTTGTACTTG 226
OY 4727 GATTGAGGAGATGATGAAGAAATTAATTAATGAGCGGAGACAGAGTAACATCACTGA 4786
DB 225 GATTGAGGAGATGATGAAGAAATTAATTAATGAGCGGAGACAGAGTAACATCACTGA 166
OY 4787 CTTAAGAGCTGAAGCTGGGTAGAGATTAGCATGCTGGAATTAATTAAGCAGCAATCAAA 4846
DB 165 CCTTGAAGCTGGGTAGAGATTAGCATGCTGGAATTAATTAAGCAGCAATCAAA 106
OY 4847 CGAAGACAGTGTCCAGCTACAGCAGTATGCAAACTTGGCAATTTTGTATTTTGG 4906
DB 105 CGAAGACAGTGTCCAGCTACAGCAGTATGCAAACTTGGCAATTTTGTATTTTGG 49

OY 4907 TATAGCTTTTAACTGTACAGCAATTCCTTATTAAGCTGTCACTATGACATTT 4966
DB 48 -----TTCAGCTGTGATGATTTGTAGTGAAGTGCATACATGACATTT 3

OY 4967 GT 4968
DB 2 GT 1

RESULT 5

BG402207 859 bp mRNA linear EST 12-MAR-2001

LOCUS 602465764F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4594023 5',

DEFINITION mRNA sequence.

ACCESSION BG402207

VERSION BG402207.1 GI:13295655

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 859)

AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Clontech Laboratories, Inc.

cDNA Library Preparation: Clontech Laboratories, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LMC134 row: 0 column: 16

High quality sequence stop: 629.

Location/Qualifiers

1. 859

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4594023"

/clone_lib="NIH_MGC_75"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: kidney; Vector: pNR-LIB (Clontech); Site_1:

SfiI (ggcgccctggc); Site_2: SfiI (ggcgccctggc); 5' and

3' adaptors were used in cloning as follows: 5' adaptor

sequence: 5'-CACGGCATTAAGGCC-3' and 3' adaptor sequence:

5'-ATCTGAGGCGGAGCGGCGGCAGTGTG-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.65

kb (range 0.5-4.0 kb). 15/15 colonies contained inserts

by PCR. This library was enriched for full-length clones

and was constructed by Clontech Laboratories (Palo Alto,

CA). Note: this is a NIH-MGC Library."

BASE COUNT 183 a 233 c 242 g 201 t

ORIGIN

Query Match

Best Local Similarity 87.8%; Score 502; DB 12; Length 859;
Pred. No. 1.2e-138;
Matches 571; Conservative 0; Mismatches 75; Indels 4; Gaps 2;

OY 3899 TCCAGGAGCAAAATTAATGAGGAGCTGGAAGAAATTAATGCGCTTAACCTGATGTGA 3958
DB 1 TCCAGGAGCAAAATTAATGAGGAGGAGCTGGAAGAAATTAATGCGCTTAACCTGATGTGA 60
OY 3959 CATCAATGAGTCCCTGGGTGCTACAGATGATCCAGAGAAATCTTTTGATCTACTGTGATAT 4018
DB 61 TGTAGAGTGTCCCTGGGTGCTACAGATGATCCAGAGAAATCTTTTGATCTACTGTGATAT 120
OY 4019 CCTCTGTGTCATCTCTTCATTTGATTTGTGGAAGCCCTCAAGTGGAGCCGGAAGAATG 4078
DB 121 CCTCTGTGTCATCTCTTCATTTGATTTGTGGAAGCCCTCAAGTGGAGCCGGAAGAATG 180
OY 4079 TCTTGAAGCATTTGTAGGGGGGTGTGGCCCAACCAATCTTGGCCCTGGCAAGTCAAG 4138

```
Db 181 TCCGTGAAGGGCTTGTAGGGGGGTGTGGCCACCCACATTCCTGGCCCTGGCAGATCAG 240
QY 4139 TCTCAGACAAAGGTTTGAAGACATCTGTGTGAGGACACCTTAATATCCAGAGTGGGT 4198
Db 241 TCTTAGACAAAGGTTTGAATGATCTGTGTGAGGACACCTTGAATATCCAGAGTGGGT 300
QY 4199 GCTGATGCTGCTCATGCTGTGTGAAGATCCTCAAGGCTTCACTCAAAAGTCACTCCT 4258
Db 301 GTTGACTGCTGGCCACTGCTGTGAGAAATGCCCAAGGCTTCATCCCTCAAGGTCATCCT 360
QY 4259 GGGTGCACACCAAGAGTGAACCTCGAATCTCATGTTCCAGAAATAGAGTGTCTAGGCT 4318
Db 361 GGGTGCACACCAAGAGTGAATCTCGAACCGCATGTTCCAGAAATAGAGTGTCTAGGCT 420
QY 4319 GTTCTTGGAGCCCAACACAGCATATTTGCTTGTCTAAAGCTCAGACGCTGCGCTCAT 4378
Db 421 GTTCTTGGAGCCCAACACAGAAATATTTGCTTGTCTAAAGCTCAGACGCTGCGCTCAT 480
QY 4379 CACTGACAAAGTAAATGACACCTGTCTGCGCATCCCGACACTACATGTCACGCGCCAGAC 4438
Db 481 CACTGACAAAGTAAATGACACCTGTCTGCGCATCCCGACACTACATGTCGCTGACCGGAC 540
QY 4439 TGAATGTTACATCACT--GGCTGGGGAAGAAACCCAAAGTACCTTTGGGA--CTGCGCTTC 4494
Db 541 CGAATGTTTCATCACTGCTGTGGGGAAGAAACCCAAAGTACCTTTGGGAGCTGCGCTTTC 600
QY 4495 TCAAGGAAGCCGCTCCTCTGTTATGAGATGAATGATGTCATCACTAT 4544
Db 601 TCAAGGAAGCCGCTCCTCTGTTATGAGATGAATGATGTCATCACTAT 650

RESULT 6
AV695776 788 bp mRNA linear EST 16-JAN-2002
LOCUS AV695776 GKC Homo sapiens cDNA clone GKCGXB08 5', mRNA sequence.
DEFINITION AV695776
ACCESSION AV695776.1 GI:10297639
VERSION EST.
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 788)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgchgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Source
1..788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GKCGXB08"
/clone_1lb="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); site_1: EcoRI; site_2:
XhoI"

BASE COUNT 196 a 195 c 227 g 167 t 3 others
```

```
ORIGIN
Query Match 9.28; Score 461.4; DB 10; Length 788;
Best Local Similarity 85.5%; Pred. No. 1,8e-126;
Matches 538; Conservative 0; Mismatches 86; Indels 5; Gaps 2;

QY 3584 CCCGAAAAATATACCAATATGATGGCTGACATGATCTACTGACAGAAATCCAGATCCGA 3643
Db 2 CCCGAAAAATATACCAATATGATGGCTGACATGATCTACTGACAGAAATCCAGATCCGA 61
QY 3644 TACAGG--CCCTTGGTGTTTTACATGAGACCCCAACATCAGGTGGAGTACTGCAACCTGA 3702
Db 62 TAAAGGCCCCCTGGGTGTATTACACAGAACCCCAACGCTCAGAGTGGAGTACTGCAACCTGA 121
QY 3703 CGCATGCTCAGACAGAAAGGAGCTGGTGGCTCTCCCTCCGACTGTCATCAGAGTTCGCA 3762
Db 122 AAAATGCTCAGAGACAGAGCGAGTGTGTAGACCTCCGCTGTCTCTGCTTCAG 181
QY 3763 GCTTAGGCTCTCTTCTGAAACAGACTGATGTTTGGAAATGGGAAAGATATACCGGGCA 3822
Db 182 ATGTAGAGACTCTCTCCGAAGAGACTGTATGTTGGAAATGGGAAAGATATACCGAGCA 241
QY 3823 AGAAGGCAACCACTGTCTACTGAGACGCGCATGCGCAAGATGGGCTGCCAGAGCCCATTA 3882
Db 242 AGAGGCGACCACTGTACTGAGACGCGCATGCGCAAGACTGGGCTGCCAGAGCCCATTA 301
QY 3883 GACACAGACGTTCAATTCACAGCAAAATAATGGCAGGCTGGAATAAATTACTGCG 3942
Db 302 GACACAGCATTTTCACTCAGACAGCAAAATCCAGCGGCGGGCTGTGAAAAAATTACTGCG 361
QY 3943 GTAACTGTATGTGATCAATCAATGATGTCCTGCTCTACACAAATCCAAACTTTT 4002
Db 362 GTAACTGTATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 417
QY 4003 TTGACTAGTGTATATCCCTCTGCTGCTGATCTGATCTGATGATGATGATGATGATGATGAT 4062
Db 418 ACGACTAGTGTATATCCCTCTGCTGCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 477
QY 4063 TGGAGCCGAGAAATGTCTCTGAGAGCATTTAGAGGGGTGTGTGGCCACCCATTCCT 4122
Db 478 TGGAGCCGAGAAATGTCTCTGAGAGCATTTAGAGGGGTGTGTGGCCACCCATTCCT 537
QY 4123 GGCCCTGGCAAGTCAAGTCTCAAGACAGTTTGGAAAGCACTTGTGTGGAGGCACTTAA 4182
Db 538 GGCCCTGGCAAGTCAAGTCTTGAACAGAGTTTGAATGCACTTGTGTGAGAGCACTTGAT 597
QY 4183 TATCCCAAGAGTGGGTGCTGACTGCTGCT 4211
Db 598 ATCCCAAGAGTGGGTGAGACTGCTGCACT 626

RESULT 7
A1061613 622 bp mRNA linear EST 11-NOV-1999
LOCUS A1061613/c
DEFINITION HA0380 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
sequence.
ACCESSION A1061613
VERSION A1061613.1 GI:6358910
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 622)
AUTHORS Yu,Y., Zhang,C., Luo,L., Ouyang,S., Li,W., Wu,J., Zhou,S., Liu,M.
and He,F.
Expression profile analysis of a human fetal liver cDNA library
unpublished (1998)
COMMENT Contact: Yongtao Yu
Department of Hematology
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Tel: 0086-10-68159479
```

FEATURES
source

1..622

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Human fetal liver cDNA library"

/tissue_type="liver"

/dev_stage="fetal"

/lab_host="MC1061/P3"

/note="Vector: pCDNA1"

BASE COUNT 167 a 167 c 131 g 157 t

ORIGIN

Query Match 8.8%; Score 441.4; DB 9; Length 622;

Best Local Similarity 84.4%; Pred. No. 1.5e-120;

Matches 539; Conservative 0; Mismatches 56; Indels 44; Gaps 2;

4372 CCGTCATCATCTGACAAAGTAATGCGACCTGTCTGCCATCCCGACACTACATGCTACCG 4431

622 CGGTCACTGACGACAAAGTAATGCGACCTGTCTGCCATCCCGAAATTAATGTGCTGCG 563

4432 CCAAGACTGATGTATCATCTGCTGGGGGAGAAACCAAGTACCTTTGGAGCTGCC 4491

562 ACCGACCGAATGTTTATCATCTGCTGGGGGAGAAACCAAGTACCTTTGGAGCTGCC 503

4492 TTTCAGAGAGCCGACCTCTTGTATTTGAGATGATGATGATGATGATGATGATGATG 4548

502 TTTCAGAGAGCCGACCTCTTGTATTTGAGATGATGATGATGATGATGATGATGATG 443

4549 -----ATATTTGCTGACATTTGGCCAGAGGACATGACA 4584

442 TGATGGAAGAGTCCATCCACGACCTCTGCTGGGCACTTTGGCCGAGGACATGACA 383

4585 GTTCCAGGCTGACATGAGGCGCTCTGCTGCTTCCGAGAGCAAAATACATTTAC 4644

382 GTTCCAGGCTGACATGAGGCGCTCTGCTGCTTCCGAGAGCAAAATACATTTAC 323

4645 AAGGACACTTCTTGGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4704

322 AAGGACACTTCTTGGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 263

4705 GTGTTTAAAGGTTTGTACTGATGAGGAAATGATGAGAAATTAATTAATGAGCGGA 4764

262 GTGTTTAAAGGTTTGTACTGATGAGGAAATGATGAGAAATTAATTAATGAGCGGA 203

4765 GACAGATGAGCACTCACTTACCTTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 4824

202 GACAGATGAGCACTCACTTACCTTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 143

4825 GAAATTAATGAGCACTCACTTACCTTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 4884

142 GAAATTAATGAGCACTCACTTACCTTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 83

4885 TGGCATTTTGTATTTTGTATTAAGCTTTTAAGCTGACGCAAAATCTGTATTA 4944

82 CGGATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTA 40

4945 AGGTGCTATGATGATGATTTTAAATAAATAAATCTG 4983

39 AGGTGCTATGATGATGATTTTAAATAAATAAATCTG 1

RESULT 8

LOCUS BG428018 762 bp mRNA linear EST 14-MAR-2001

DEFINITION 602501424F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4615157 5',

ACCESSION BG428018 mRNA sequence.

VERSION BG428018.1 GI:13334524

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 762)

NIH-MGC <http://imgc.ncl.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://imgc1367.row.p.columb.06>

Plate: LICM1367 row: P column: 06

High quality sequence stop: 737.

location/Qualifiers

1..762

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="IMAGE:4615157"

/clone_lib="NIH_MGC_75"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: kidney; Vector: pBMR-LIB (Clontech); Site:1;

5' adaptor sequence: 5'-CACGGCCATTATGAGC-3' and 3' adaptor

sequence: 5'-ATCTAGAGCCGAGGCGGCGACATG-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.65

kb (range 0.5-4.0 kb). 15/15 colonies contained inserts

by PCR. This library was enriched for full-length clones

and was constructed by Clontech Laboratories (Palo Alto,

CA). Note: this is a NIH-MGC library."

BASE COUNT 192 a 160 c 203 g 207 t

ORIGIN

Query Match 7.9%; Score 396; DB 12; Length 762;

Best Local Similarity 82.5%; Pred. No. 7.6e-107;

Matches 525; Conservative 0; Mismatches 80; Indels 31; Gaps 5;

4349 CTGCTTAAGCTTAAGCAGGCTGCGCTATCATCTGACAAAGTAAATGCACTGCTGCC 4408

2 CTGCTTAAGCTTAAGCAGGCTGCGCTATCATCTGACAAAGTAAATGCACTGCTGCC 61

4409 ATCCCAAGTACATGATGACACCCGAGTGAATGATGATGATGATGATGATGATGATG 4468

62 ATCCCAAGTACATGATGACACCCGAGTGAATGATGATGATGATGATGATGATGATGATG 121

4469 CCAAGTACCTTGGAGCTGCGCTTCTCAAGAAAGCCGACCTCTGTTATTTGAGATGA 4528

122 CCAAGTACCTTGGAGCTGCGCTTCTCAAGAAAGCCGACCTCTGTTATTTGAGATGA 181

4529 AGTGGCAATCATATAAGT-----ATATTTGCTGTA 4561

182 AGTGGCAATCATATAAGT-----ATATTTGCTGTA 241

4562 GCATTGGCCAGAGGACCTGACCTGACAGGCTGACAGTGAAGGCGCTGCTGTTGCTT 4621

242 GCATTGGCCAGAGGACCTGACCTGACAGGCTGACAGTGAAGGCGCTGCTGTTGCTT 300

4622 CGAGAGGACAAATATTAATTTACAGAGTCACTTCTGGGGCTTGGGGCTGACAGGCC 4681

301 CGAGAGGACAAATATTAATTTACAGAGTCACTTCTGGGGCTTGGGGCTGACAGGCC 360

4682 CAATTAAGCTGCTGCTATGCTGCTGTTCAAGTTT-TGTTACTTGATTTGAGGAATGA 4740

361 CAATTAAGCTGCTGCTATGCTGCTGTTCAAGTTT-TGTTACTTGATTTGAGGAATGA 420

4741 TGAGAAATTAATTAAT-TGAGAGGAGACAGAGTGAAGCACTCACTCACTTAAGAGTGA 4799

421 TGAGAAATTAATTAATTTGAGAGGAGACAGAGTGAAGCACTCACTCACTTAAGAGTGA 480

[illegible]

Db	184	ATGTAGAGACTCTCTTCCGAGAGAGCTGTATGTTGGGATGGGAAAGGAGATACCGAGGC	243
QY	3822	AAAGAGGCAACCACTGTACTGTGGACGCCATGTCCAGSATAAGGCTGCCAGAGGCCAT	3881
Db	244	AAAGAGGCGGACCACTGTACTGTGGACGCCATGTCCAGSATAAGGCTGCCAGAGGCCAT	303
QY	3882	AGACACAGCAGCGCTATTCACAGGCAAAATMAATGGGCGACGTCTGGAAAAAAATTTACTGC	3941
Db	304	AGACACAGCATTTTACTCTCCAGAGCAAAATCACAGGCGGGGTCTG-----AAATTTACTGC	359
QY	3942	CGTACCCCTGATGTGTACATCAATGTCCTGTGTCTTACACAAATGAATCCAGAAAACTT	4001
Db	360	CGTACCCCTGATGTGTATGTAGTGTGTCTCCGTGTCTTACACAAATGCA-----	411
QY	4002	TTTGACTACTGTGTATATCCCTCTGTGTCAATCCCTTCATTGATGTGGAAAGCCTCA	4061
Db	412	-----GTGGAAAGCCTCA	425
QY	4062	GTGGAGCCGAGAAATGTCTTGGAGCATTTGAGGGGGGTGTGTGGCCACCCATTTCC	4121
Db	426	GTGGAGCCGAGAAATGTCTTGGAGGGGTGTGTGGGGGGGTGTGTGGACCCACCATTTCC	485
QY	4122	TGGCCCTGGCAGTACTGTCTGAGAACAAAGTTTGGAAAGCACTTCTGTGAGGCACCTTA	4181
Db	486	TGGCCCTGGCAGTACTGTCTGAGAACAAAGTTTGGAAATGACATCTGTGTGAGGCACCTTT	545
QY	4182	ATATCCCGAGAGGGGTGTGTACTCTGTCTGACATGCTTGAAGAAAGTCCTGAAGCCTTCA	4241
Db	546	ATATCCCGAGAGGGGTGTGTACTCTGTCTGACATGCTTGAAGAAAGTCCTGAAGCCTTCA	603
QY	4242	TCCTACAGGTCATCTGGGTGCACA	4267
Db	604	ATCTACAGGTCATCTGGGTGCACA	629
RESULT 10	BE348267/c	593 bp	EST 18-JUL-2000
LOCUS	BE348267	593 bp	mRNA
DEFINITION	hw21e11.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3183596 3'		
ACCESSION	BE348267		
VERSION	BE348267.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 593)		
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaabs-r@mail.nih.gov		
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.		
	Emmert-Buck, M.D., Ph.D.		
	cDNA Library Preparation: M. Bento Soares, Ph.D.		
	DNA Sequencing by: Greg Lennon, Ph.D.		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL, send email to:		
	infoimage.llnl.gov		
	Seq primer: -40up from Gibco		
	High quality sequence: 425.		
FEATURES	Location/Qualifiers		
SOURCE	1..593		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3183596"		
	/clone_lib="NCI_CGAP_Kid11"		
	/lab_host="DH10B"		
	/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with		

a modified polylinker. Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 132376-132911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 161 a 158 c 122 g 152 t
ORIGIN

Query Match 7.8%; Score 392; DB 10; Length 593;
Best Local Similarity 82.6%; Pred. No. 9.9e-106;
Matches 504; Conservative 0; Mismatches 60; Indels 46; Gaps 3;

OY 4411 CCCGAGCTACATGCTACCCGAGAGCTGAATGTACATCATGCTGGGAGAAACC 4470
||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 593 CCCCAATTAATGCTGCTGACCGGACCAATGTTTCATCATGCTGGGAGAAACC 534
OY 4471 AAGTACCTTGGGAGCTGCTCTCAAGAAACCCAGCTCCTGTATATGAGATGAG 4530
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 533 AAGGTAC -TTTGAGCTGCTCTCAAGAAACCCAGCTCCTGTATGAGATTAAG 476
OY 4531 TGTCAATCACTATAGT-----ATATTGCTGTAGC 4563
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 475 TGTCAATCGCTATGAGTTCCTGATGAAAGACTCCATCCACCGAATCTGTCTGGGC 416
OY 4564 ATTGGCCAGAGCACTGAAGATTGCCAGGCTGACAGTGGAGGCTCTGGTTCTTCG 4623
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 415 ATTGGCCGAGCACTGACAGTGGCCAGGCTGACAGTGGAGGCTCTGGTTCTTCG 356
OY 4624 AGAAGCAATAATACATTTTACAAGAGTCACTTGGGGCTTGGCTGTGACGCCCA 4683
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 355 AGAAGCAATAATACATTTTACAAGAGTCACTTGGGGCTTGGCTGTGACGCCCA 296
OY 4684 ATAAGCTGCTGTATGCTGCTGCTTTCAAGGTTTGTACTTGGATGAGGAATGATGA 4743
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 295 ATAAGCTGCTGTATGCTGCTGCTTTCAAGGTTTGTACTTGGATGAGGAATGATGA 236
OY 4744 GAATTAATTAATGAGGAGAGAGAGTGAACATCAACTACTTGAAGCTGAACGT 4803
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 235 GAATTAATTAATGAGGAGAGAGAGTGAACATCAACTACTTGAAGCTGAACGT 176
OY 4804 GGGTAAGATTAGCATGCTGGAATTAATGAGCAATCAACGAGCACTGTGCCA 4863
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 175 GGGTAAGATTAGCATGCTGGAATTAATGAGCAATCAACGAGCACTGTGCCA 116
OY 4864 GCTACAGCTATGCAAAACCTTGGCATTTTGTATTGTAATGAAGCTTTAAGTTC 4923
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 115 GCTACAGCTATGCAAAACCTTGGCATTTTGTATTGTAATGAAGCTTTAAGTTC 73
OY 4924 TGACTGCAAAATTTGTATTAAGTGTCAATAGTATGACATTTGTTAAATTAACCTGTG 4983
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 72 TGACTGCTGATTCGTATGTAAGGTGACATAGTATGTAATTAATTAACCTGTG 13
OY 4984 CACTTATTTT 4993
||||| ||
DB 12 TACTTAATT 3

RESULT 11
BI219735 701 bp mRNA linear EST 11-JUL-2001
LOCUS 602936156F1 NCI_CGAP_Li9 Mus musculus cDNA IMAGE:5099748 5',
DEFINITION mRNA sequence.
ACCESSION BI219735
VERSION BI219735.1 GI:14673179
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 701)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cqabds-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHM11239 row: 0 column: 13
High quality sequence stop: 699.
Location/Qualifiers

FEATURES
source

1. 701
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5099748"
/clone_lib="NCI_CGAP_Li9"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo df. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 183 a 189 c 200 g 129 t

ORIGIN

Query Match 7.7%; Score 387.4; DB 13; Length 701;
Best Local Similarity 75.9%; Pred. No. 2.7e-104;
Matches 532; Conservative 0; Mismatches 161; Indels 8; Gaps 4;

OY 3495 CAGAGTATCGAGGACATTCACACACAGTGTGACAGAGAGCATGTAATCTGTCA 3554
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 CAGAGTATCGAGGACATTCACACACAGTGTGACAGAGAGCATGTAATCTGTCA 61
OY 3555 TCCATGACACACACACGCGCATCAGAGACCCAGAAAATACCAATGATGGCTGACA 3614
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 GGTATGTTTCCACACAGCATTTCCAGACAGACCCACAGAAATGCTGCTGGAG 121
OY 3615 ATGAACTACTGAGGAATCCAGATGCGATTCAGAGCTTGGTGTTCATGAGACCC 3674
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 122 ATGAACTACTGAGGAATCCAGATGCGATTCAGAGCTTGGTGTTCATGAGACCC 181
OY 3675 AGCATCAGGTGGAGTACTGCAACCTGACGGATGCTCAGACACAGAGGAGCTGTGTC 3734
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 182 AGCATCAGGTGGAGTACTGCAACCTGACGGATGCTCAGACACAGAGGAGCTGTGTC 241
OY 3735 GCTCCCGACGTGATCAGAGTTCAGAGCTTGGAGGCTCTCTGTAACAAGCTGTATG 3794
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 242 GAATTCGCCACAGTTCCTCAGGAACCAATGCGCGGACGACATCTGACACATCTCATG 301
OY 3795 TTTGGGAATGGGAAGATACCGGGGCAAGAAGCAACACTGTACTGGAGCCCATG 3854
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 302 TATGGGAATGGGAAGATACCGGGGCAAGAAGCAACACTGTACTGGAGCCCATG 361
OY 3855 CAGGAATGGGCTGCCAGAGACCCCATGACACAGCAAGCTTCATCCAGGACCAATAA 3914
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 362 CAGGAGTGGGCTGCCAGAGACCCCATGACAGCAAGCTTCATCCAGGACCAATAA 421
OY 3915 TGGCAGAGTCTGGAATAAATTAATGAGCTTAACCTGATGATGATGATGATGATG 3974
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 422 CCGGAGAGTCTGGAATAAATTAATGAGCTTAACCTGATGATGATGATGATGATG 481
OY 3975 TGTACAC -AATGAATCCAGAAAATTTTGTACTGATGATGATGATGATGATG 4032
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 482 TGTATACAAAAAACCACAGAAAATTTATGATATTTGATGATGATGATGATGATG 541
OY 4033 -CTCTTCATTTGATTTGGGAAGCTCAAGTGGAGCGAAGAAATGCTGGAGAT 4090
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 542 AGCATCATTCTTGTGAGTGGGAAACCTCAGGTGAACCAAGAAATGCTGGAGGAT 601

QY 4091 TGTAGGGGGGTGTGTGGCCACCACCAATTCGTGGCCGTGGGACAGCACTGTCCACAAAG 4150
 Db 602 GGTGGGTGGCTGGCTGGT-G-ACAACTTCACCTCTGGGCCCTTGGCAATACAGGCTTAGACAAAG 660
 QY 4151 GTTT--GGAAGACACTCTGTGTGGAGGACACCTTAATATCCC 4188
 Db 661 ATTATCCGGACAGACACTTCTGTGGCCGTACTTTAATAGGCC 701

RESULT 12	
AV662084	
LOCUS	
DEFINITION	AV662084 GLC Homo sapiens cDNA clone GICHA06 3', mRNA sequence.
	611 bp mRNA linear EST 16-JAN-2002

VERSION	AV662084.1	GI:9883098
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 611)
 Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.
 AUTHORS

TITLE
Xiao, H., Qiu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W., Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X., Hu, G., Gu, J., Chen, Z., and Han, Z.
In-situ insight into hepatocellular carcinogenesis at transcriptome level

JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE	21625106
COMMENT	Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai
331 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex. 45)
Fax: 86-21-50801922
Email: hanzy@chgc.sh.cn
This clone is available at CHGC in Shanghai.

```

FEATURES
source
location/Qualifiers
1..611
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCHAG06"
/clone_id="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2
XhoI"

```

BASE COUNT	144 a	150 c	164 g	153 t
ORIGIN				
Query Match		7.7%;	Score 383.4;	DB 10;
Best Local Similarity		83.6%;	Pred. No. 3.9e-103;	Length 611;

Matches	516;	Conservative	0;	Mismatches	71;	Indels	30;	Gaps	6;
Oy	4090	TTGTATGGGGGGGTGTGGGCCACCCACCAATTCCTGGGCGCGCAAGTCAGTCCAGACAA							4149
Db	2	TTGTATGGGGGGGTGTGGGCCACCCACA-TCTTGCGCTTGGCAATGTCAGTCTTAACACA							60

QY	4150	GGTTTGGAAAGCACTTCTGTGGAGGCACTTAATATATCCCAAGTGGGTGACTGCTG	4209
Db	61	GGTTTGGAAATGCACTTCTGTGGAGGCACTTAATATCCCAAGTGGGTGACTGCTG	120

QY 4330 CCACGCAAGCAGATATTGGCTTGTCTAAAGACTAAGAGCGCTTCGCATCACTGACGAAG 4389
| | | | |
Db 241 CCACGCAAAAAATATTTCCTTGCTTTAAAGACTAAGAGTCTCCCTCATCACTGACGAAG 300

QY 4390 TATATCCAGACTTGATGCCATCCCCAGACTACATGATGACCGCACAGATAATTACA 4449
| | | | |
Db 301 TATATCCAGACTTGATGCCATCCCCAATTAATGTGTGCGCGACCGCAAGCAATTTTCA 360

QY 4450 TCACGTGGCTGGGGAGAAAACCAGAAGTACCTTTGGAGCTGGCCCTTCTCAAGAACCCAGC 4509
| | | | |
Db 361 TCACGTGGCTGGGGAGAAAACCAGAAGTACTTTTGGAGCTGGCCCTCTTAAGAAAGCC - AC 418

QY 4510 TCCCTGTATTAGAAATGAAGTGTGCATCACTATAAGI ----- 4548
| | | | |
Db 419 TCCCTGTATTAGAAATGAAGTGTGCATCACTATAAGI ----- 477

QY 4549 --ATATTGTGCTGAGCAATTTGGCCAGAGGCACTGACATTTGCCAGGGTGACAGTGAGG 4606
| | | | |
Db 478 ACCGAACTTTGGGTGGGCAATTTGCCGAGAGCACTGACAGTTG -CAAAGTGAAGTGAGG 536

QY 4607 GCCTGTGGTTGCTTCGAGAAAGACAAATATCAATTTTACAGGAGTCACTCTTTGGGGTCT 4666
| | | | |
Db 537 CCGTTTGGTTTGT--GAGAAAGCAAAATCAATTTTACAGGAGTCCCTTTTGGGGTCT 594

QY 4667 TGGCTGTGACGCCCCA 4683
| | | | |
Db 595 TGggTgcccccccccA 611

RESULT 13			
LOCUS	H60805	481 bp	mRNA linear
DEFINITION	EST 06-OCT-1995		
	Yr5g90.r1	Soares fetal liver spleen	INTLS Homo sapiens cDNA clone
	IMAGE:208288.5	' similar to gb.X05199	PLASMINOGEN PRECURSOR (HUMAN
); mRNA sequence.		

ACCESSION	H60805
VERSION	H60805.1
KEYWORDS	GI:1013637
SOURCE	EST.
ORGANISM	human.
	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Euharipoda; Mesozoa; Chordata; Ctenaria; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnaria; Hominidae; Homo. 1 (bases 1 to 481)	Hallier, L., Clark, N., Dubuque, T., Ellison, K., Hawkins, M., Holman, R., Hultman, M., Kucaba, T., Le, M., Lennon, G., Mairra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, P., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.	The WashU-Merck EST Project	Unpublished (1995)	Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
High quality sequence stops: 339
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: M13RP1
High quality sequence stop: 339.

```

FEATURES
source
Location/Qualifiers
1..481
/organism="Homo sapiens"
/db_xref="GDB:3777419"
/db_xref="taxon:9606"
/clone="IMAGE:208288"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;"

```


ACCESSION BG616686
VERSION BG616686.1 GI:13668057
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 758)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs@email.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LUCM1595 row: 1 column: 24
High quality sequence stop: 553.
Location/Qualifiers
1..758
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4733687"
/clone_lib="NIH-MGC_76"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
sfi1 (ggccgcctggcc); Site_2: sfi1 (ggccatattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATATGAGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.85
kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH-MGC Library."
BASE COUNT 242 a 150 c 206 g 160 t
ORIGIN
Query Match 7.3%; Score 365; DB 12; Length 758;
Best Local Similarity 82.7%; Pred. No. 1.6e-97;
Matches 459; Conservative 0; Mismatches 55; Indels 41; Gaps 2;
QY 4462 GAGAAACCCAGGTAAGCTTTGGGACTGGCTTCTCAAGAGCCAGCTCTTTGTTATG 4521
|||||
Db 1 GAGAAACCCAGGTAAGCTTTGGGACTGGCTTCTCAAGAGCCAGCTCTTTGTTATG 60
QY 4522 AGAATGAAGTGTCAATCACTTAAGT-----ATATTT 4554
|||||
Db 61 AGAATGAAGTGTCAATCACTTAAGT-----ATATTT 120
QY 4555 GTTGCTAGCATTTGGCAGAGGCTGACAGTGGAGGGGCTCTGG 4614
|||||
Db 121 GTTGCTAGCATTTGGCAGAGGCTGACAGTGGAGGGGCTCTGG 180
QY 4615 TTTGCTTCGAGAGCAAAATACATTTTACAAGGAGTCATTCTTGGGGTCTTGGCTGTG 4674
|||||
Db 181 TTTGCTTCGAGAGCAAAATACATTTTACAAGGAGTCATTCTTGGGGTCTTGGCTGTG 240
QY 4675 CACGCCCAATAAGCTGTGTCTATGCTCGTGTTCAGAGTTTGTACTTGGATTGAGG 4734
|||||
Db 241 CACGCCCAATAAGCTGTGTCTATGCTCGTGTTCAGAGTTTGTACTTGGATTGAGG 300
QY 4735 GAATGATGAGAAATATTAATTTGAGCGGAGACAGAGTGAACATCAACCTACTTAGAG 4794
|||||
Db 301 GAATGATGAGAAATATTAATTTGAGCGGAGACAGAGTGAACATCAACCTACTTAGAG 360
QY 4795 CTGAAACGTGGGTAGGATTTAGCATGCTGGAATATATATAGACAGCATCAAGAGACA 4854
|||||
Db 361 CTGGAACGTGGGTAGGATTTAGCATGCTGGAATATATATAGACAGCATCAAGAGACA 420

QY 4855 CTGTCCAGCTACACGCTATGCAAAACCTTGGCATTTTGGTATTTTGTGATTAAGCT 4914
|||||
Db 421 CTGTCCAGCTACACGCTATGCAAAACCTTGGCATTTTGGTATTTTGTGATTAAGCT 480
QY 4915 TTTAAGTCTGAGTACAAATTCGTATTAAGGTGTCAATAGCTATGACATTTGTTAAAAA 4974
|||||
Db 481 T-----GGATTCGTGTAGTAAAGGTGACATAGCTATGACATTTGTTAAAAA 526
QY 4975 TAAACTTGCACCTTA 4989
|||||
Db 527 TAAACTTGCACCTTA 541

Search completed: March 6, 2003, 15:20:05
Job time: 6051.5 secs

